UTS UNIVERSITY OF TECHNOLOGY SYDNEY

Investigating genetic and molecular factors contributing to sexual dimorphism in lung disease

by Karosham Diren Reddy

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of:
Distinguished Professor Brian Oliver
Dr Alen Faiz
Dr Yik Lung Chan

University of Technology Sydney
Faculty of Science

November 2023

Certificate of Original Authorship

I, Karosham Diren Reddy, declare that this thesis is submitted in fulfillment of the requirements for the award of Doctor of Philosophy in the School of Life Science, Faculty of Science at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note: Signature removed prior to publication.

Date: 08/11/2023

Acknowledgements

My experience as a PhD candidate will be a time I will cherish and positively reflect on for the rest of my life. I have been fortunate to meet, work and socialise with amazing people who have enriched my studies.

My deepest thanks go to Distinguished Professor Brian Oliver. Routinely, I would enter Brian's office questioning my ability, feeling dejected. Inevitably, I'd always leave energised and motivated to keep going. He always recognised and encouraged my ability, even when I did not. I have inherited parts of Brian's humour and sense of curiosity, which I hope I can carry throughout my life.

Thank you to Dr Alen Faiz. I am immensely grateful for your time, patience and willingness to explain complex bioinformatic methods and concepts. I could never have generated such a detailed and technical body of work without Alen's teachings.

I want to recognise my collaborators from Groningen, The Netherlands and Newcastle, Australia. I could not have completed this work without your insight and expertise. I'm also grateful to all the tissue donors, study participants, and teams who collected the human samples used in this thesis.

I'm blessed to have worked with all past and present members of the Cell Biology group, the RBMB group and the Woolcock Institute of Medical Research. A special thanks to Dr Dia Xenaki, from the beginning, your ability to know when I needed help and your willingness to answer my silly questions kept me moving forward. Thank you for always checking on me daily and encouraging me not to eat lunch at my desk so I wouldn't lose my mind.

To Raz, Jack, Sandra, Roy, Katrina, and Jeremy, thank you for showing me the ropes and being wonderful mentors and great friends. You were all unique and extraordinary examples of the work ethic and persistence needed to complete a PhD. Thank you to Maria, Patrick, Senani and Ollie for sharing and enduring all the trials and tribulations. Thank you, Akshay and Anna, who I turned to and relied upon many times. Andra and Kamila, thank you for ensuring I was (excessively) well-fed and providing endless entertainment.

Thanks to my friends for always being my biggest supporters, forgiving my absences, and always being there to help me relax or lift my spirits after long and tiresome days.

Finally, my immeasurable gratitude goes to my mum, dad and sisters. You have provided me with the foundation to always strive for my best and chase anything I aspire towards. I hope you can share my sense of achievement and recognise your time, energy and effort in this thesis. Thank you for ensuring I ate healthily and always had a fresh cup of tea.

Format of Thesis Statement

I hereby state that this is a submission of a thesis by compilation.

Posters, presentations and awards throughout this candidature

Oral Presentations

1. **"Sex cells: X and Y chromosome genes in asthma"**; <u>**Reddy, K.D.**</u> (2023); Invited speaker: Woolcock Institute of Medical Research Seminar Series.

2. **"Sex Cells: Imbalance in X and Y chromosome"**, <u>**Reddy, K.D.**</u>, Rathnayake, S.N.H., Faiz, A. & Oliver, B.G.G. (2022); Rapid Fire presentation at the National Asthma Meeting.

3. **"Sex Cells: Imbalance in X and Y chromosome-linked genes"**, <u>**Reddy, K.D.**</u>, Rathnayake, S.N.H., Faiz, A. & Oliver, B.G.G. (2022); Rapid Fire Presentation at the Australian Society for Medical Research Annual Scientific Meeting 2022.

4. **"Sex Cells: The contribution of the X and Y chromosome in asthma"**, <u>**Reddy, K.D.**</u>, Rathnayake, S.N.H., Faiz, A. & Oliver, B.G.G. (2022); Presentation at the Thoracic Society of Australia and New Zealand NSW/ACT Branch Update meeting 2022.

5. **"Investigating the factors causing sexual dimorphism in lung disease"**, Reddy K.D. (2021); Invited speaker: Woolcock Institute of Medical Research Seminar Series

6. **"Novel functions of X and Y chromosome-linked genes in lung disease"**, <u>**Reddy, K.D.**</u>, Rathnayake, S.N.H., Faiz, A. & Oliver, B.G.G. (2021); Oral presentation at the Woolcock Institute of Medical Research Postgraduate Symposium.

7. **"Let's talk about sex – Don't forget the Y chromosome"**, <u>**Reddy, K.D.**</u>, Rathnayake, S.N.H., Faiz, A. & Oliver, B.G.G. (2021); Oral Presentation as part of the 3-minute thesis competition, University of Technology Sydney.

8. **"Investigating sex differences in lung disease"**, <u>**Reddy, K.D.**</u>, Rathnayake, S.N.H., Faiz, A. & Oliver, B.G.G. (2021); Oral Presentation at the Early Career Researcher Think Tank (Woolcock Institute, Heart Research Institute & Centenary Institute).

9. **"Smoking alters gene expression and methylation patterns in asthma patient nasal epithelium"**, <u>**Reddy K.D.**</u>, Lan A., Boudewijn I.M., Rathnayake S.N.R., Koppelman G.H., Oliver B.G.G., van den Berge M. & Faiz A (2021); Oral Presentation at European Respiratory Society Lung Science Conference 2021.

10. **"Sexual Dimorphism in respiratory disease"**, <u>**Reddy, K.D.**</u>, Rathnayake, S.N.H., Faiz, A. & Oliver, B.G.G. (2020); Oral Presentation at St Vincent Hospital Research Week.

11. **"Epigenetic-like changes to signalling proteins in COPD"**, <u>**Reddy, K.D.**</u>, Zakarya, R., Xenaki, X. & Oliver, B.G.G. (2019); Oral Presentation at New Horizons 2019.

Poster Presentations

1. **"The Y-chromosome regulates hallmark features of asthma and COPD"**, <u>**Reddy, K.D.**</u>, Rathnayake, S.N.H., Xenaki, D., Faiz, A. & Oliver, B.G.G. (2022); Poster presentation at European Respiratory Society International Congress 2022.

2. **"Investigating the effect of *in utero* exposure to PM and cigarette smoke on asthma in offspring"**, <u>**Reddy, K.D.**</u>, Chan, Y.L., Allam, V.S.R.R., Faiz, A., Chen, H. & Oliver, B.G.G. (2022) Poster presentation at the Thoracic Society of Australia and New Zealand Annual Scientific Meeting 2022.

3. **"Current smoking affects gene expression and methylation patterns in asthma patient nasal epithelium"**, <u>**Reddy K.D.**</u>, Lan A., Boudewijn I.M., Rathnayake S.N.R., Koppelman G.H., Oliver B.G.G., van den Berge M. & Faiz A (2021); Poster Presentation at the American Thoracic Society International Conference 2021.

4. **"Smoking changes gene expression and methylation in asthmatic nasal epithelium"**, <u>**Reddy K.D.**</u>, Lan A., Boudewijn I.M., Rathnayake S.N.R., Koppelman G.H., Oliver B.G.G., van den Berge M. & Faiz A (2021); Poster presentation at the Thoracic Society of Australia and New Zealand Annual Scientific Meeting 2021.

5. **"Epigenetic regulation of signalling proteins in COPD"**, <u>**Reddy, K.D.**</u>, Zakarya, R., Xenaki, X. & Oliver, B.G.G. (2019); Poster Presentation at European Respiratory Society International Congress 2019.

Awards

1. National Asthma Meeting; Awarded **'Best Poster Pitch' (2022)**, based on the merit of the presentation

2. Thoracic Society of Australia and New Zealand; Awarded **'Best abstract presentation on asthma' (2022)**, based on the merit of the presentation

3. University of Technology Sydney; **Vice-Chancellor's conference travel scholarship**, based on the quality of the application

4. University of Technology Sydney – **Runner-up in 3-minute thesis competition (2021)**; based on the merit of the presentation

5. Early Career Researcher Think Tank; Awarded **'Best Presentation by a PhD student' (2021)**, based on the merit of the presentation

6. Woolcock Institute of Medical Research; Awarded **'Best PhD student presentation' (2021)**, based on the merit of the presentation

7. European Respiratory Society; Awarded a **bursary to attend the Lung Science Conference (2021)** based on the merit of the submitted abstract

8. New Horizons; Awarded **'Best presentation by a young investigator' (2019)**, based on the merit of the presentation

9. Thoracic Society of Australia and New Zealand; Awarded **'Best abstract presentation on airway disease' (2019)**, based on the merit of the presentation

10. Lung Foundation Australia – **A Menarini Australia Travel Grant (2019)**; Award to attend European Respiratory Society International Congress 2019 to present an abstract in the form of a poster.

Publications generated from this work

The following publications contribute directly to the understanding and arguments presented in this thesis.

1. **'Sexually Dimorphic Production of Interleukin-6 in Respiratory Disease'** (2020), *Physiological Reports*; **Reddy K.D.**, Rutting S., Tonga K., Xenaki D., Simpson J.L., McDonald V.M., Plit M., Malouf M., Zakarya R. & Oliver B.G.G

2. **'Sex specific effects of in utero and adult tobacco smoke exposure'** (2021), *American Journal of Physiology – Lung Cellular and Molecular Physiology*; **Reddy, K.D**. & Oliver B.G.G.

3. **'Current smoking alters gene expression and DNA methylation in the nasal epithelium of patients with asthma'** (2021), *American Journal of Cell and Molecular Biology*; **Reddy K.D.**, Lan A., Boudewijn I.M., Rathnayake S.N.R., Koppelman G.H., Oliver B.G.G., van den Berge M. & Faiz A.

4. **'Sexual dimorphism in chronic respiratory diseases'** (2023), *Cell & Bioscience*; **Reddy, K.D.** & Oliver B.G.G.

Publications produced adjunct to the current work

The publications listed below do not directly contribute to the current work but are included as evidence of research productivity and skill development.

1. **'Immunological Axis of Berberine in Managing Inflammation Underlying Chronic Respiratory Inflammatory Disease'** (2020) *Chemico-biological interactions*; Tew N.X, Lau, N.J.X., Chellappan D.K., Madheswaran T., Zeeshan F., Tambuwala M.M., Aljabali A.A.A., Balusamy S.R., Perumalsamy H., Gupta G., Oliver B.G.G., Hsu A., Wark P., **Reddy K.D.**, Wadhwa R., Hansbro PM. & Dua K.

2. **'BET proteins are associated with the induction of small airway fibrosis in COPD'** (2021), *Thorax*; Zakarya R., Chan Y.L., Rutting S., **Reddy K.D.**, Bozier J., Woldhuis R.R., Xenaki D., Van Ly D., Chen H., Brandsma C., Adcock I.M. & Oliver B.G.G.

3. **'Airway Smooth Muscle cells from severe asthma patients with fixed airflow obstruction are responsive to steroid and bronchodilator treatment in vitro'** (2021), *ERJ Open Research*; Rutting S., Xenaki D., **Reddy K.D.**, Baraket M., Chapman D.G., King G.G., Oliver B.G.G. & Tonga, K.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

A549	Adenocarcinoma human alveolar basal epithelial cells
AGC	Automatic gain control
AGRF	Australian genomes research facility
AHR	Airway hyperresponsiveness
ANOVA	Analysis of variance
APS	Ammonium persulfate
ARACNe	**A**lgorithm for the **R**econstruction of **G**ene **R**egulatory **Ne**tworks
ASM	Airway smooth muscle
BCA	Bicinchoninic acid
BH	Benjamini-Hochberg
BM	Basement membrane
β-ME	Beta-mercaptoethanol
BMI	Body mass index
BSA	Bovine serum albumin
cDNA	Complementary DNA
CD8+ T cell	Cytotoxic T cell
CHORDC1	Cysteine and histidine rich domain containing 1
CO$_2$	Carbon dioxide
Col4α1	Collagen 4 alpha chain 1
COPD	Chronic Obstructive Pulmonary Disease
CPM	Counts per million
CRISPR	Clustered regularly interspaced short palindromic repeats
CSE	Cigarette smoke extract
CXCL8	Chemokine (C-X-C motif) ligand 8
DEG	Differentially expressed genes
DGE	Differential gene expression
DKO	Double-knockout of UTX and UTY cell line
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ECM	Extracellular matrix
EGR	Early growth response 1
ELISA	Enzyme-linked immunosorbent assay
Ep	Epithelial
ERT	Estrogen replacement therapy
ETDA	Ethylenediaminetetraacetic acid
FAO	Fixed airflow obstruction
FBS	Fetal bovine serum
FC	Fold-change
FC	Fold change
FCG	Four core genome
FDR	False discovery rate
FEV$_1$	Forced expiratory volume in 1 second
FEV$_1$ % Pred	Percentage of the patient FEV$_1$ divided by the average FEV$_1$ of the population
FN1	Fibronectin
FOXOA1	Forkhead box protein A1

FVC	Forced vital capacity
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GFP	Green fluorescent protein
GOLD	Global initiative for chronic obstructive pulmonary disease
gRNA	guide RNA
GSVA	Gene set variation analysis
H3K27	Lysine residue 27 on histone H3
H3K27me3	Trimethylation of Lysine-27 residue on histone H3
hASM	Human airway smooth muscle
HBSS	Hanks balanced salt solution
HCD	Higher-energy C-trap dissociation
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HREC	Human research ethics committee
HRT	Hormone replacement therapy
HRP	Horseradish peroxidase
HRV	Human rhinovirus
HSP90ab1	Heat shock protein 90 alpha family class B
IAA	Iodoacetamide
ICS	Inhaled corticosteroids
IgE	Immunoglobulin E
IgG	Immunoglobulin G
IL	Interleukin
IL-1β	Interleukin – 1 beta
indel	Insertion or deletion of DNA bases
ITGA4	Integrin subunit alpha 4
ITGB3	Integrin subunit beta 3
ITGB8	Integrin subunit beta 8
kDa	kilodalton
KDM6A	Lysine Demethylase 6A
KDM6C	Lysine Demethylase 6C
KO	Knockout
LB	Lysogeny broth
LC-MS/MS	Liquid chromatography - tandem mass spectrometry
LFQ	Label free quantification
mCOPD	mild COPD
MES	2-(N-morpholino) ethane sulfonic acid
mRNA	messenger ribonucleic acid
MSY	Male-specific region of the Y-chromosome
MTT	3-4, 5- dimethylthiazol-2-(yl)-,5, diphenyltetrazolium
NaCl	Sodium chloride
NADPH	nicotinamide adenine dinucleotide phosphate
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NH$_2$OH	Hydroxylamine
NIH	National institute of health
NKRF	NF-κB repressing factor
NK	Natural killer
Nlp	Ninein-like protein
NSCLC	Non-small cell lung cancer

OD	Optical density
OLiVIA	Effects **O**f Extra-fine Partic**L**e HFA-Beclomethasone **V**ersus Coarse Particle Treatment **I**n Smokers and Ex-smokers with **A**sthma
Opti-MEM	Optimal - minimal essential medium
PAM	Protospacer adjacent motif
PBS	Phosphate buffered saline
PC20	Methacholine provocation concentration provoking 20% decrease in FEV1
PCA	Principle component analysis
PCR	Polymerase chain reaction
PEFR	Peak expiratory flow rate
PMSF	Phenylmethylsulphonyl fluoride
pQTL	Protein quantitative trait loci
PVDF	Polyvinylidene difluoride
RIN	RNA integrity number
RNA	Ribonucleic acid
RNA-seq	RNA-sequencing
ROS	Reactive oxygen species
RP	Ribosomal protein
rpm	revolutions per minute
RPS4X	Ribosomal Protein S4 X-linked
RPS4Y1	Ribosomal Protein S4 Y-linked 1
rRNA	Ribosomal RNA
RSV	Respiratory syncytial virus
SCLC	small cell lung cancer
sCOPD	severe COPD
SD	Standard deviation
SDB-RPS	Styrene divinylbenzene-reverse phase sulfonated
SDS-PAGE	Sodium dodecyl sulfate - polyacrylamide gel electrophoresis
SEM	standard error of the mean
SEO-COPD	Severe early-onset COPD
SHERLOCk	An integrative genomic approach to **S**olve t**H**e puzzle of sev**ER**e ear**L**y-**O**nset **C**OPD
siRNA	small interfering RNA
SPE	Solid phase extraction
STAT3	Signal transducer and activator of transcription 3
TCEP	tris (2-carboxyethyl) phosphine
TEMED	N, N, N', N'-tetramethyl ethylenediamine
TGF-β1	Transforming growth factor beta
TMB	3, 3', 5, 5' -tetramethylbenzidine
TNC	Tenascin C
TNFα	Tumour necrosis factor alpha
T-PBS	Tween20 – phosphate buffered saline
T-TBS	Tween20 – tris buffered saline
UMAP	Uniform manifold approximation and projection for dimension reduction
UTX	Ubiquitously transcribed tetratricopeptide repeat, X chromosome
UTY	Ubiquitously transcribed tetratricopeptide repeat, Y chromosome
X-chr	X chromosome
XCi	X-chromosome inactivation

Y-chr Y chromosome
ZFX Zinc Finger Protein X-linked
ZFY Zinc Finger Protein Y-linked

Abstract

Differences in disease susceptibility, progression and severity patterns exist between males and females. Asthma and chronic obstructive pulmonary disease (COPD) are pervasive respiratory diseases that demonstrate sexual dimorphism. Male children experience increased asthma rates and worse health outcomes than female children, with this trend reversing after puberty. Importantly, sex differences exist for the hallmark features of asthma and COPD in inflammation, airway fibrosis and remodelling, and regulation of cell death. Despite a clear relationship between biological sex and disease, the contributory genetic and molecular factors remain poorly understood. Preliminary analysis of patient-derived primary airway cells identified that female-derived cells produce increased levels of the proinflammatory cytokine IL6 compared to male cells, irrespective of disease diagnosis. Therefore, sex differences in immunoregulation are observable *ex vivo*, where biological factors such as sex hormones are removed. Thus, intrinsic molecular and genetic factors such as the sex chromosomes likely contribute to this difference.

Highly similar but non-exact gene homologs on the X and Y chromosomes demonstrate imbalanced expression. Females express double the X-chromosome-linked homolog, whilst only males express the Y-chromosome-linked version. This imbalanced expression or function of X and Y chromosome homologs may contribute to sex differences in inflammatory, fibrotic and cell death-related processes. We aimed to explore, characterise and compare the function of the *ZFX/ZFY, RPS4Y1/RPS4X* and *UTX/UTY* gene pairs that have vital genome regulatory functions. This includes mediating gene transcription, protein translation and histone demethylation. As such, they may specifically regulate pathological processes between males and females.

CRISPR-Cas9 knockout cell lines were generated for each candidate gene in male-derived A549 cells. Disease-relevant phenotypes were analysed in knockout cells, such as immunoregulation, cell proliferation, cell adhesion, cell death and extracellular matrix protein production. RNA-sequencing and proteomics analyses were paired with independent patient cohorts to identify pathways regulated by the candidate genes and highlight whether they contribute to clinically relevant outcomes, such as lung function measurements.

For the first time, we show novel differences in the function of the XY gene pairs that directly relate to the hallmark disease features of inflammation, fibrosis and cell death. These divergent functions may contribute to sex differences in the susceptibility and severity of asthma and COPD. We relate gene pathways regulated by the candidate genes to clinical measurements. This vital data provides a foundation for identifying target pathways for the development of new, more effective treatments to improve patient outcomes.

Chapter 1 – Introduction

1.1 Publication Declaration - 'Sexual dimorphism in chronic respiratory diseases'

Karosham D. Reddy & **Brian G.G. Oliver** (2023) *Cell & Bioscience*.

Status: Published.

Author Contributions: Both authors conceived the idea, read the manuscript and edited the final version. KDR drafted the manuscript. Both authors read and approved the final manuscript.

Signatures:

Name	Signature	Date
Karosham D. Reddy	Production Note: Signature removed prior to publication.	**01/03/2023**
Brian G.G. Oliver	Production Note: Signature removed prior to publication.	**01/03/2023**

REVIEW

Open Access

Sexual dimorphism in chronic respiratory diseases

Karosham Diren Reddy[1,2]* and Brian Gregory George Oliver[1,2]

Abstract

Sex differences in susceptibility, severity, and progression are prevalent for various diseases in multiple organ systems. This phenomenon is particularly apparent in respiratory diseases. Asthma demonstrates an age-dependent pattern of sexual dimorphism. However, marked differences between males and females exist in other pervasive conditions such as chronic obstructive pulmonary disease (COPD) and lung cancer. The sex hormones estrogen and testosterone are commonly considered the primary factors causing sexual dimorphism in disease. However, how they contribute to differences in disease onset between males and females remains undefined. The sex chromosomes are an under-investigated fundamental form of sexual dimorphism. Recent studies highlight key X and Y-chromosome-linked genes that regulate vital cell processes and can contribute to disease-relevant mechanisms. This review summarises patterns of sex differences in asthma, COPD and lung cancer, highlighting physiological mechanisms causing the observed dimorphism. We also describe the role of the sex hormones and present candidate genes on the sex chromosomes as potential factors contributing to sexual dimorphism in disease.

Keywords Sexual dimorphism, Asthma, COPD, Lung cancer, Sex chromosomes, Inflammation, Remodelling

Background

Sexual dimorphism refers to a divergence in the physical characteristics between chromosomally defined males and females of a species. These differences exist at the organ, cellular and molecular levels and are critical for establishing differences between males and females and enabling sexual reproduction [1–3]. However, the potential for these differences to contribute to sex differences in susceptibility and disease development is overlooked. Recently, the scientific community has actively aimed to recognise and investigate trends of sex differences both epidemiologically and physiologically. For example, the National Institutes of Health (NIH) mandated that sex

must be considered a critical biological variable [4]. This instruction highlights the lack of data investigating sex as a biological factor in disease development and progression. The NIH highlights that sex is a biological variable that should be considered at all levels of research, from experimental design to analysis and reporting findings in animal and human studies.

It is critical to define the difference between biological sex and gender. Biological sex refers to the sex chromosome complement of an individual. Males carry one Y-chromosome, and one X-chromosome (XY), whilst females have two X-chromosomes (XX). The presence of the *SRY* gene on the Y-chromosome initiates a hormone cascade during early development, stimulating the formation of the characteristic male phenotype. In contrast, the absence of the *SRY* gene results in the generation of female characteristics. Gender is defined by social norms and expectations for how "men and women" should behave [5]. The factors that influence gender vary between different cultures and with time. Notably, a growing body of work recognises the complex interaction

*Correspondence:
Karosham Diren Reddy
Karosham.Reddy@sydney.edu.au
[1] Respiratory and Cellular Molecular Biology Group, Woolcock Institute of Medical Research, Glebe, NSW 2037, Australia
[2] School of Life Science, University of Technology Sydney, Ultimo, NSW 2007, Australia

between gender and disease outcomes. Although important, the impact of gender on disease is beyond the scope of the current body of work, and biological sex differences between males and females will be the primary focus of this review.

The reporting and investigation of sex differences in disease are being increasingly recognised across various health conditions [6]. Nonetheless, there remains an incomplete understanding of the molecular and genetic factors driving sexual dimorphism. This is partly a result of large clinical and cohort studies designating sex as a confounding factor or a covariate in the data analyses [7]. As a result, the complexities of diseases remain poorly understood or unidentified as sometimes the effects of disease between males and females may occur in opposing directions, resulting in a "net-zero" effect size when grouped [8]. When public RNA-seq datasets are stratified by sex, significant differences in gene expression are apparent between males and females in non-gonadal tissues, which are otherwise non-significant when unstratified [8, 9]. As a result, a considerable gap exists in our understanding of the fundamental differences between males and females. Sex differences in response to the same clinical interventions are well reported in the literature to affect patient outcomes [7, 10–12]. Developing a deeper understanding of the fundamental factors and mechanisms driving sexual dimorphism in diseases is critical to furthering our understanding of disease development and creating new, more effective 'personalised' clinical treatments.

Here, we will review patterns of sex differences in prominent respiratory diseases and present how sexual dimorphisms manifest at a molecular and physiological level. We will also explore how the sex hormones and sex chromosomes contribute to pathological differences between males and females.

Sexual dimorphism in lung physiology

Differences in the lung structure between males and females may contribute to patterns of sexual dimorphism in various respiratory diseases. The lung's development and maturation present a complex and dynamic pattern of sexual dimorphism driven by various factors. Importantly, differences in lung physiology between males and females have important clinical implications. Male lungs are bigger than female lungs, with this difference existing from birth into adulthood [13]. The disparity in lung development in utero between male and female foetuses begins as early as 16 to 24 weeks gestation [14]. Female foetuses have smaller airways and a lower number of respiratory bronchioles compared to males; however, their maturation rate is faster. Surfactant, an essential compound enabling correct lung function [15], is produced

earlier in females than males, enabling a faster lung maturation rate. The faster rate of development is thought to explain why female neonates are less likely to suffer from respiratory distress syndrome compared to male neonates [16]. Estrogen produced by the placenta stimulates the production of surfactant and the development of alveoli [17]. In contrast, testicular-derived androgens such as testosterone function to suppress the production of surfactant [13, 17], to which female foetuses are not exposed. As a result of divergent patterns of lung development, in early life, males and females present with distinct physiological lung profiles. As mentioned above, female lungs are smaller, with fewer respiratory bronchioles and smaller airways [14], whilst the luminal area for the large and central airways is approximately 14–31% larger in males, even when matching for lung size [18]. Cumulatively, as the female lung is smaller, with fewer respiratory bronchioles, the total number of alveoli and lung surface area is higher for males throughout early development. This disproportionate lung size and airway growth rate is called 'dysanapsis' [19, 20]. Females demonstrate higher forced expiratory flow rates until they are 18 years old [21]. This increased airflow rate is postulated to reduce female children's susceptibility to damage due to in utero exposures and the development of respiratory conditions such as asthma and respiratory tract infections [6, 22]. As total lung capacity (TLC) increases in females, sex differences in expiratory flow rates also diminish. Clinical studies attribute sex differences in lung pathology changes such as airway fibrosis and inflammation to physiological and anatomical differences [23]. As such, biological and structural differences between males and females may contribute to patterns of sexual dimorphism in respiratory diseases such as asthma, chronic obstructive pulmonary disease (COPD) and lung cancer.

Sex differences in respiratory diseases

As mentioned, sexual dimorphism is apparent in a range of diseases across multiple organ systems. However, sex differences in the susceptibility, severity and progression between males and females for chronic respiratory diseases are particularly intricate. For example, idiopathic pulmonary fibrosis is two times more common in males [24], whilst cystic fibrosis demonstrates greater severity in women [25]. This complex interaction between sex and disease becomes significantly apparent for asthma, chronic obstructive pulmonary disease (COPD) and lung cancer.

Asthma

Asthma is a heterogenous respiratory disease characterised by hyper-reactive and reversible airway inflammation. It is primarily diagnosed based on a history of

respiratory symptoms from wheezing, episodic shortness of breath (dyspnea), chest tightness and cough, varying over time [26]. Intrinsic to asthma is a complex interplay between airway inflammation and remodelling, culminating in airway hyperresponsiveness (AHR). Airway remodelling refers to structural changes in the airways, such as increased airway smooth muscle (ASM), thickened basement membrane, epithelial dysplasia, and increased collagen deposition [27]. These changes result in a thickened airway wall, which, combined with inflammatory exudate produced by immune cells, obstruct the airway causing difficulty breathing. Figure 1 illustrates the significant obstruction of the airway lumen in asthma patients compared to healthy patients. The extent of expiratory airflow limitation is measured by the forced expiratory volume in one second (FEV_1), which is a common tool used to evaluate asthma. Variability in FEV_1 is commonly triggered by exposures such as exercise, allergens or viral infections. In most cases, as asthma becomes more severe the airflow obstruction becomes fixed, increasing the rate of FEV_1 loss.

The exact cause of asthma remains unknown. As such, it remains a significant health problem. Approximately 300 million people suffer asthma worldwide, with Australia demonstrating one of the highest prevalence rates, at 11.2% [28, 29]. Asthma is the leading disease burden for children younger than 15, with children who are under 15 more likely to be hospitalised with asthma than those older than 15 [30]. Part of the difficulty in treating and managing asthma relates to its heterogeneous nature, with five primary clinical phenotypes recognised. A range of clinical patterns or treatment responses determines the different asthma phenotypes. These include the causative agents (allergic), timing of diagnosis

(adult-onset), lung function outcome (persistent airflow limitation) and associated comorbidities [26]. Allergic asthma is the most common phenotype, often starting in childhood and driven by eosinophilic airway inflammation in response to stimulation by allergens. Non-allergic asthma, on the other hand, presents a more neutrophilic immune cell profile and responds less to inhaled corticosteroids (ICS). Significant advances have been made regarding asthma management and treatment; however, the exact cause of asthma remains elusive. Furthermore, modern treatments are flawed, with therapies such as targeted monoclonal antibodies failing to eliminate dangerous exacerbations [31]. Other clinical interventions only target specific aspects of asthma, such as anti-inflammatory steroids that reduce the immune response but do not resolve the structural changes that occur [31]. This inability to resolve airway remodelling limits the ability to control patients with severe disease and presents a significant shortfall in the attempt to cure/reverse asthma.

Sex differences in asthma

Compounding the complexity of asthma is its sexually dimorphic nature. Differences between males and females with asthma exist in childhood and adulthood asthma. Young males (12.1%) report a higher prevalence of asthma compared to young females (7.9%) and are two times more likely to be hospitalised due to asthma [32]. However, this pattern reverses after puberty, towards an increased asthma diagnosis in adult females (13.9%) compared to adult males (9.6%), as depicted in Fig. 2. However, some conjecture exists in the literature, with some studies concluding no sex differences in asthma severity [33, 34]. Other studies report higher mortality and a higher rate of exacerbations in females [23], severely
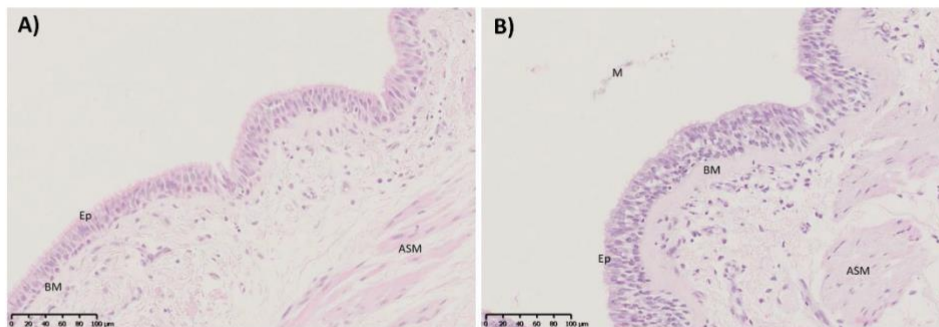
Fig. 1 Micrographs of airways from a healthy patient **A** and an asthmatic patient **B** stained with haematoxylin and eosin. A thickened basement membrane (BM) can be seen in the asthmatic patient with hyperplasia of the epithelial (Ep) layer. A noticeable increase in the airway smooth muscle (ASM) thickness can also be seen. A slight mucus exudate can be seen in the airway lumen of the asthma patient. Scale bars = 100 μm
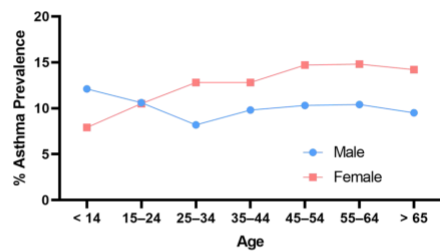
Fig. 2 Asthma prevalence by age group and sex in Australia recorded by the National Health Survey 2017–2018 [29]. The pink line indicates female asthma prevalence (%), and the blue line indicates male asthma prevalence (%)

reducing their daily quality of life. In contrast, adult females have a three-fold increased risk for hospitalisation due to asthma [32]. Several factors contribute to the sexual dimorphic presentation of asthma, including anatomical differences, sex hormones and environmental/occupational factors.

The apparent differences in asthma susceptibility and severity between the sexes are associated with many factors. The differential inflammatory, fibrotic, and remodelling processes present a complex nexus of physiological and molecular factors contributing to asthma pathogenesis and progression. The sex hormones have been linked with the temporal shift in asthma susceptibility [35–37]. However, genome studies reveal genetic associations in asthma differ by sex, highlighting unique biological underlying factors contributing to sexual dimorphism in asthma [8]. Network analyses conclude that asthma may function via differential mechanisms in males or females, despite involving similar processes and functional outcomes [38]. Therefore, a holistic understanding of environmental, genetic, hormonal and physiological factors is needed to unravel the complex interaction between asthma and biological sex.

Sex differences in inflammation and remodelling in asthma

A predominance of CD4+ and Th2 cells with infiltration of eosinophils and mast cells in the airways characterises type 2 airway inflammation. CD4+ and T helper cells initiate and perpetuate the phenotype of prolonged inflammation [39]. This type of inflammation primarily drives allergic/atopic asthma. In particular, it has been noted that young males present with more allergic inflammation [40]. Interestingly, males are less prone to immunological illnesses over their lifetime than females [23]. This pattern has been linked to the role of sex hormones, which will be discussed in detail later. Males and females demonstrate distinct immune cell populations in asthma.

Female lungs have increased levels of type 2 lymphoid cells, specifically a subset of cells that do not express killer-cell lectin-like receptor G1—which is absent from male lungs [41]. After puberty, this cell population can produce type 2 inflammatory cytokines, thus creating different pro-inflammatory environments between males and females with asthma. There is a distinct pattern towards more atopic asthma, airway infections and bronchiolitis in young males before puberty, with more males admitted to hospital before puberty [39]. The differences between male and female immunological mechanisms and responses are complex and change with age.

Tumour necrosis factor—alpha (TNFα) and transformation growth factor—beta (TGFβ) are prominent immunoregulatory cytokines closely associated with asthma pathogenesis [42]. They are critical to asthma's cellular and humoral immune responses [43]. Associations have been identified between genetic polymorphisms in atopic and non-atopic asthma patients [44, 45]. In particular, they have been correlated with serum IgE levels, of which boys demonstrate higher levels than girls [28]. However, the allergic response to asthma allergens can depend on CD4+ and CD8+ cells rather than IgE levels [23]. As such, the CD4+ to CD8+ cell ratio can be a marker of chronic lung disease. This ratio is lower in males than females in adulthood [46], contributing to sex differences in the inflammatory response. TNFα demonstrates various pathological functions, including inducing the infiltration and activation of immune cells through promoting increased expression of adhesion molecules [47], thus increasing bronchial hyperresponsiveness. Further, TNFα dysregulates epithelial barrier activities along with IL-13 [48]. Female-biased expression of IL-13 in asthma patients may interrupt tight junction proteins, contributing to worse asthma symptoms in females. The IL-17 pathway is also up-regulated in females with asthma, potentially driving increased airway hyperresponsiveness [49]. Type 2 immune response in asthma increases the number of neutrophils driven by CXCL8, a well-known neutrophil chemotactic cytokine [23, 50]. Similarly, IL6 levels correlate with worse asthma outcomes as part of type 2 immune response causing neutrophil infiltration [51]. Neutrophilic asthma is linked with a poorer response to corticosteroids [10], impacting patient outcomes. Furthermore, asthmatic males demonstrate reduced response to β2-agonists with age despite treatment with inhaled steroids [27].

Airway remodelling refers to the degradation and repair of the ECM and the increased proliferation of fibroblasts. Traditionally it is believed that inflammation drives the airway remodelling in asthma, progressing into AHR and culminating in fixed airflow obstruction (FAO) [52]. However, there is a growing consensus that

the altered structure of the airway may stimulate and promote inflammatory processes [53, 54]. For example, the breakdown of collagen IV affects asthma severity [45, 46] due to a decrease in the tumstatin fragment, reducing inflammation and AHR [55]. Therefore, sex differences in inflammation may drive sex differences in fibrosis and vice versa. Rasmussen et al. [27] found in a longitudinal population study that airway remodelling is associated with the male sex, with reduced lung function outcomes from childhood into adulthood. Males demonstrate an accelerated decline in FEV_1 predicted values, potentially driven by higher rates of fixed airflow obstruction in younger and older populations [28, 56]. Despite this, female mice exposed to an Ova-sensitised model experience significantly more airway remodelling than male mice [57]. Lung function has been used as a surrogate method to measure the progression of airway remodelling and asthma, as an increased rate of FEV_1 decline is seen in many asthma cases [52]. The ratio of FEV_1 to vital capacity (the total volume of air that can be inhaled) indicates a downward trend in females from late adolescence into adulthood, signifying greater fibrotic and remodeling changes compared to males [27]. In early development, young males demonstrate slower expiratory airflow rates despite having similar total lung volumes to young females [39], indicating an initial structural disadvantage. Therefore, young males are more liable to develop asthma symptoms at a younger age, which might contribute to worse outcomes.

A complex relationship exists between inflammation, airway remodelling, biological sex and asthma. The exact mechanisms and factors causing these apparent differences in pathological processes remain unclear and require deeper investigation and discussion. Structural or functional sex differences are unlikely to drive respiratory disorders such as asthma and wheezing, with hormonal or genetic factors likely to contribute.

Chronic obstructive pulmonary disease (COPD)

COPD is characterised by progressive and irreversible airflow limitation, culminating in a sustained decline in lung function. Chronic inflammation of the airway drives the thickening and narrowing of the airway structural layers, obstructing airflow [58], similar to asthma. In contrast, COPD is characterised by small airway remodelling and emphysema, with occlusion of the airways and parenchymal destruction. COPD patients experience significant airflow limitation which is presented through the hallmark features of chronic cough, shortness of breath and excessive mucous production [59]. The irreversible worsening of disease symptoms significantly reduces patient quality of life, eventuating in disability and death. COPD is currently the third leading cause of

death worldwide, affecting 7.5% of Australians older than 40 and 30% of people older than 75 [60]. Cigarette smoking is the best-known risk factor for the development of COPD; however, the exact cause of COPD remains unknown, with only 10% of COPD cases attributable to genetic factors. It is generally considered an adult-onset disease, and lifelong exposure to environmental factors functions as a critical pathogenic factor.

Sex differences in COPD

Historically, COPD was considered male predominant. In recent years, there has been increased awareness and investigation of sex differences in COPD incidence and health outcomes. The National Centre for Health Services (NCHS) data shows that COPD death rates declined for males yet remained steady in females over time [61]. This trend is driven by a normalisation of smoking rates between males and females, which has narrowed COPD diagnosis rates between the sexes. Female smokers are 50% more likely to develop COPD than males [62]. Tam et al. found that 60% of all COPD hospitalisations occurred in females [63]. When considering the effect of smoking further, females demonstrate worse lung function and disease prognosis than males despite smoking at the same level [64] (Fig. 3). As such, due to environmental exposures, females present a steeper decline in lung functionality, contributing to increased rates of COPD diagnosis [65]. Females showed a 5.7% reduction in FEV_1% predicted in the low smoke exposure group compared to males [66]. Women report more dyspnoea, chronic cough and lower overall quality of life scores. In addition, females also represent a more significant proportion
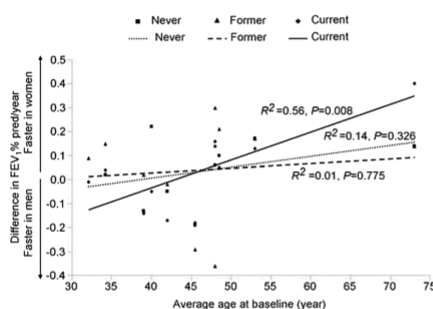
Fig. 3 Unweighted analysis of the relationship between age and gender-related differences in the annual decline in lung function (FEV_1%) pred/year according to smoking. The squares/dotted line represents never smokers, the triangles/dashed line represents former smokers, and the diamonds/solid line represents current smokers. *Recreated with permission from BioMed Central publisher and was first published by Gan* et al. [64]

of severe early-onset COPD patients (SEO-COPD) [67], defined by the development of COPD before 50 years of age with lung function less than the lower limit of normal [62, 68]. Importantly, sex differences in COPD manifest at a pathophysiological level, with males demonstrating higher levels of emphysema, whilst females show more significant small airway disease. A mouse model by Tam et al. found female mice exhibit increased levels of small airway remodelling and greater activation of TGFβ in the small airways after chronic cigarette exposure [63], which was not observed for males. A proteomic investigation of bronchoalveolar lavage (BAL) revealed an increase in macrophage autophagy in females who had developed COPD [69]. Macrophage autophagy is known to be a critical process in the pathophysiology of COPD [70]. An autoimmune profile has been reported in COPD [71]. As autoimmune diseases demonstrate a distinct female-bias, this provides support of a molecular basis towards increased susceptibility of females to develop COPD [72]. Forsslund et al. show female smokers with COPD present with increased CD8 T-cells expressing CCR5 compared to non-smoker females with COPD [73]. The authors highlight distinct T-cell profiles dependent on smoking status, demonstrating a correlation between Th1 inflammation with goblet cell density and BAL macrophages in female smokers with COPD. Comparatively, a correlation between Th2 inflammation and IgG serum concentration was reported in male smokers, with no effect of COPD observed. As such, clinical observations of sex differences stem from distinct cellular and biological interactions with cigarette smoking and COPD development.

Lung cancer

Lung cancer is one of the most common cancer types, with rates continuing to increase globally [74, 75]. Overall, there is a trend towards increased lung cancer cases in never-smoking individuals, although 80% of all lung cancer cases are attributable to a history of tobacco smoking [76]. Lung cancer is broadly classified into two subgroups: non-small cell lung cancer (NSCLC), which comprises 85% of cases, and small cell lung cancer (SCLC), which accounts for 15% of patients [77]. NSCLC includes specific subtypes such as squamous cell carcinoma, large cell carcinoma from epithelial cells that line the bronchus, and adenocarcinoma from the gland tissues [78]. SCLC is characterised by a rapid doubling time and is the most aggressive, reporting a 5 year survival rate of less than 7% [79]. Significant genetic diversity in lung cancer complicates the investigation and understanding of biological pathways involved in disease development and progression. Advancements in modern sequencing technologies have identified key oncogenic targets such as *KRAS, EGFR, BRAF* and *JAK2* [80].

However, the complexity of lung cancer is attributed to its lack of recurrent mutations that occur at a high frequency. This phenomenon impedes the ability to develop more effect treatments [81]. Improving our knowledge of the fundamental pathological features of lung cancer will enable the identification of key, targetable pathways and ultimately improve patient outcomes.

Sex differences in lung cancer

Sex bias in lung cancer is well established, with notable differences observed since 1996 [82]. Lung cancer is the second most diagnosed malignancy and the leading cause of cancer death worldwide [77]. Females demonstrate more adenocarcinoma and less squamous cell carcinoma than males [77]. This pattern was thought to relate to differences in smoking patterns between the sexes. However, 50% of women diagnosed with lung cancer are never smokers, compared to 20% of males [83]. Thun et al. found female never smokers of European, African American and Asian descent all showed increased lung cancer rates compared to their male counterparts [84]. The combination of these trends spurred the notion that female lung cancers have a distinct genetic and pathogenic profile compared to males.

Figure 4 highlights the complex and dynamic pattern of sex differences in lung cancer over time. Male incidence of diagnosis remains steady (Fig. 4A), and shows decreased mortality rate (Fig. 4B). In comparison, females have demonstrated increasing incidence and mortality rates over the last 40 years. Interestingly, females demonstrate higher survival rates for all histological subtypes of lung cancer [85] after accounting for the stage at diagnosis, age and treatment (represented in Fig. 4C) [76]. Notably, females tend to be diagnosed with lung cancer at a younger age, potentially enabling better opportunities for treatment before disease progression [86]. An Australian longitudinal cohort study by Yu et al. found that although women have higher diagnosis rates, males demonstrate a 43% increased risk of lung cancer mortality [11, 87]. The authors identified female patients were significantly more responsive to treatments, which is supported by multiple other studies potentially contributing to increased survival rates [76, 77], whilst other studies identified sex-specific benefits depending on the treatment [88, 89]. Yu et al. identified an increased density of B-cells in the adenocarcinoma tissues of females, which they suggested might contribute to improved treatment and survival outcomes, as B cells have critical anti-tumour activity. Furthermore, as female patients tend to be younger, they also show better baseline health status compared to male counterparts at a similar stage of disease [87].

The innate and adaptive immune responses between males and females differ significantly. Studies show that
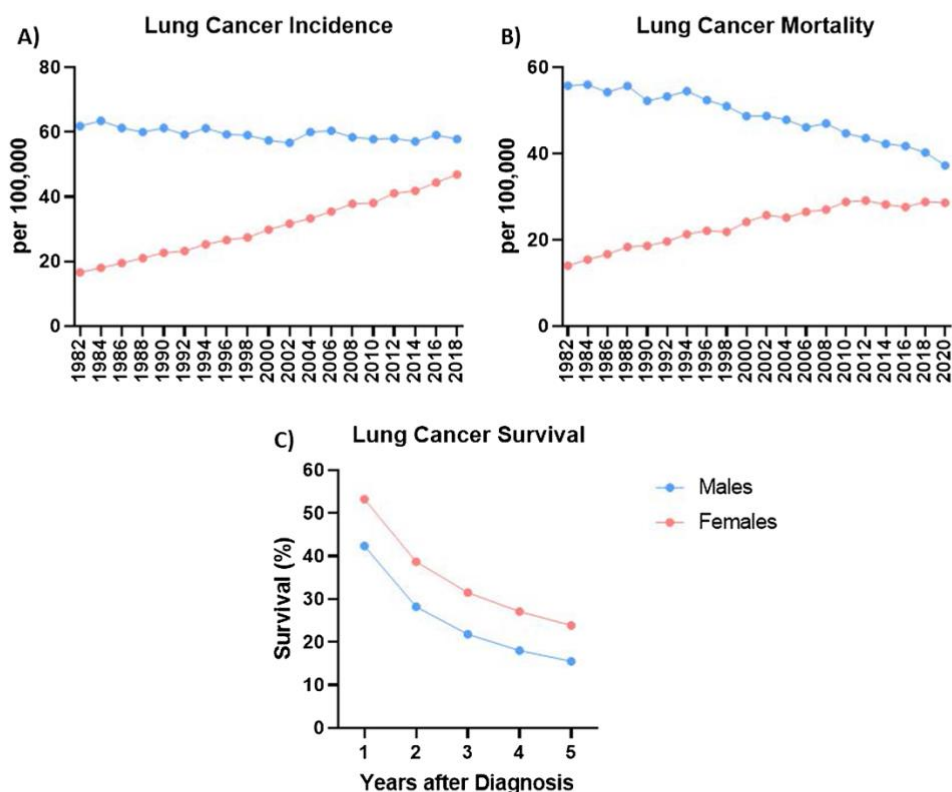
Fig. 4 Lung cancer incidence, mortality and survival, collected by the Australian Institute of Health and Welfare [90]. **A** Age-adjusted lung cancer incidence rates by sex **B** age-adjusted lung cancer mortality by sex **C** lung cancer survival rates (%) each year after diagnosis by sex [90]. The blue line represents males, whilst the pink line represents females

males with a suppressed immune response are more susceptible to developing malignancies and infections than females [91]. PD-1, an immune checkpoint inhibitor, is increased in female NSCLC patients compared to males, with higher expression in female CD4+T cells [92]. This pattern was postulated by Gu et al. to be associated with increasing testosterone levels in females with NSCLC, however, this mechanism requires further exploration [92]. Therefore, a complex relationship between lung cancer and biological sex exists. Although females are diagnosed with lung cancer more often, there is a clear bias towards worse outcomes in males. Furthermore, a distinct histological profile exists between the sexes, with unique pathological profiles potentially apparent between males and females. These differences may contribute to

variations in response to antitumor and immunotherapy treatments. The trend of lung cancer appears to be linked to social tobacco smoking patterns. However, there is a distinct shift towards increased non-smoking-related lung cancer development rates. This trend implies important genetic factors may promote epidemiological and pathologic differences between males and females.

The sex hormones: estrogen and testosterone
Sexual dimorphism is apparent in various diseases, especially in the lung. The most common and easily attributable factor that explains differences between males and females are the sex hormones. These steroid hormones are critical factors driving phenotypic differences between males and females. Androgens (testosterone)

8

are the primary male hormone, and estrogens (estradiol) are the predominant female hormone. However, it should be noted that both androgens and estrogens are found in females and males, but at lower physiological concentrations.

Estrogen and testosterone are commonly thought only to be produced by the sex organs (ovaries and testes). However, their effects extend to peripheral tissues in a paracrine manner [93]. As such, sex steroids are implicated in various diseases impacting multiple organ systems, from cardiovascular to neurological and respiratory [35]. Clinical and epidemiological studies have associated sex hormones in modulating lung diseases. The sex hormones can alter airway tone and modulate the inflammatory response, with estrogen driving a pro-inflammatory environment, whilst androgens are reported to have anti-inflammatory activity [94, 95]. This difference in effect potentially contributes to worse outcomes and increased incidence of complications in females suffering lung diseases associated with inflammation [35].

Elevated testosterone levels are linked with decreased asthma risk regardless of sex [96]. Higher levels of androgen receptor expression in human airway epithelium are associated with better lung function and fewer asthma symptoms [97]. This observation is potentially driven by the ability of androgens to attenuate inflammatory factors in the lung [32, 97] and, by inducing airway smooth muscle relaxation through decreased levels of cellular calcium [98]. Dijkstra et al. identified polymorphisms in estrogen receptor alpha (ERα) that are closely linked with airway hyperresponsiveness and worse lung function decline in females [99]. Even in the absence of stimulation, ERα deficient mice demonstrate significant AHR [100], highlighting a close relationship between estrogen and asthma symptoms.

Generally, estrogens are considered immunological enhancers (i.e. promote an immune response), and androgens/testosterone are immunological suppressors (i.e. reduce the immune response) [101]. It is hypothesised that increased estrogen levels enhance the T2 immune profile in asthma, causing the development and an increased frequency of exacerbations in asthma [101]. This notion is supported by evidence of peri-menstrual worsening of asthma symptoms reported in 20–40% of females with severe or difficult-to-control asthma [102, 103]. Symptoms appear worst during the mid-luteal phase (estrogen and progesterone are elevated) [104]. Many immune cells that predominate in asthma express estrogen and testosterone receptors such as eosinophils, airway smooth muscle cells and T cells [101]. As such, hormonal fluctuations throughout the menstrual cycle influence the immune response to allergic stimuli [105]. Epidemiological evidence shows that the rate of COPD incidence in females is increasing, with the death rate in females increasing since 2000 [106]. Differences in COPD pathology between males and females may be partly modulated by estrogen. Mouse models show female mice develop more airway obstruction upon chronic cigarette smoke exposure, while male mice develop more emphysema. However, ovariectomised female mice (to remove estrogenic effects) develop emphysema similar to male mice [63]. This indicates that estrogen promotes a different pathological COPD phenotype in females, contributing to worse disease outcomes. Androgens have been shown to have anti-inflammatory properties and regulate the structure and function of non-reproductive organs. Further, increased testosterone levels are linked with a decreased risk of asthma in both males and females [96]. Chiarella et al. [107] conducted an extensive review outlining the varying effect estrogen has on multiple airway cell types, highlighting the complex interaction between the sex hormones and the lung.

Androgens have been implicated in lung cancer development, with reduced androgen levels associated with decreased cell proliferation [108] and found to alter the gene expression profile in cancer cell lines [109]. Testosterone is believed to function as a promoter of tumour cell proliferation, contributing to the higher incidence and worse outcomes from cancer in men [110]. However, estrogen has also been linked to the incidence of NSCLC in females—compounding the adverse effects of cigarette smoking. Females who smoke using estrogen replacement therapy (ERT) indicated more than double an increased risk for adenocarcinoma development. Whereas those using ERT and who have never smoked showed no increased odds of developing lung cancer [78].

Puberty, pregnancy and menopause

We have discussed how sex-specific patterns of asthma incidence change in puberty. It is important to recognise that both estrogen and testosterone change during puberty and are active in both sexes. The dramatically increased estrogen production in females at puberty potentially promotes increased immune system responsiveness and airway smooth muscle contraction [101, 111]. Conversely, increased testosterone in males is likely protective, suppressing eosinophil and neutrophil inflammation in the lungs and improving airway tone [112]. Further, Bulkhi et al. found a one unit $\log_2$ increase of serum testosterone was associated with an 11% decreased risk of asthma in males and a 10% decrease in females [96]. However, no correlation between serum testosterone and current asthma was reported for patients under 12 years old. This highlights that childhood asthma is promoted by non-hormonal factors and requires further investigation.

An increasing prevalence of asthma in pregnancy has been reported overtime, from 3.7% in 1997 to 8.4% in 2001 [113], with rates as high as 12.7% in Australia in 2012 [114]. Approximately 20% of females with asthma experience increased exacerbations during pregnancy [115]. The mechanical implications due to uterus enlargement combined with hormonal changes during pregnancy cause increased asthma symptoms in pregnant females [116]. Hormonal changes occur in pregnancy to fulfil the mother's and fetus metabolic needs. As detailed earlier, estrogen and progesterone modulate the immune response, which can lead to worse asthma symptoms. Up to 40% of mothers report that changes to their asthma vary with successive pregnancies, indicating that a complex interplay of factors affects asthma in pregnancy [117].

Pregnancy with concurrent COPD or lung cancer is rare as both conditions develop later in life. Only two instances of pregnancy in patients with COPD have been reported. In one example, COPD symptoms improved with pregnancy, potentially due to the protective role of estrogen against increased bronchoconstriction [118]. The patient's condition significantly declined post-delivery, indicating that the pregnancy caused a partial reversal of COPD progression. In contrast, the other case of COPD in pregnancy [119] indicated little to no improvement, potentially due to the overall worse disease state of the patient. Limited data and studies are evaluating lung cancer's molecular and genomic features in pregnancy. However, adenocarcinoma is the most common form of lung cancer in pregnancy (80%), which may be linked to increased estrogen receptor expression in this cancer type [120, 121]. Consistent with general patterns for cancer, pathological characteristics and health outcomes for patients with lung cancer are the same irrespective of the pregnancy [120, 122].

Early menarche is closely linked with faster lung function decline and worse health outcomes later in life [13], with smoking known to induce early menopause. Menopause is characterised by a distinct reduction of progesterone and estrogen production in females, occurring around the fifth decade [123]. Generally, postmenopausal females have a significantly reduced risk of developing asthma [124]. However, females with asthma at menopause have high levels of circulating estradiol, with a dose-dependent correlation with asthma severity [125]. Asthma prevalence increases in males compared to females after 50 years of age, coincidentally when testosterone levels decrease further [126]. A recent systematic review [123] determined that the contribution of menopause to asthma remains conflicting due to sources of bias and heterogeneity in the current literature. The authors posit that it may be prudent to explore the relationship

between menopause and specific asthma phenotypes, which may lead to more insightful conclusions. Only two studies have investigated the link between menopause and COPD, with both finding no association [127, 128]. Hormone replacement therapy (HRT) increased the risk of adult-onset asthma by 49% in menopausal females in two independent cohorts [124, 129]. This highlights a complex interaction between asthma, menopause and hormone changes that requires deeper investigation. Early menopause is linked to an increased risk of lung cancer [130]; although smoking can bring forward the onset of menopause, this may primarily be a smoking effect. Alternatively, some studies have found late menopause (older than 55 years) is linked with an increased risk of lung cancer among non-smokers [131, 132]. This pattern may be caused by greater life-long exposure to estrogen, which has been linked with the development of other cancers [133]. Inconsistent definition of disease outcomes and measurements in studies investigating and associating menopause is a significant limiting factor. As a result, findings from these studies generate conflicting results. The use of clear clinical definitions and the examination of disease subtypes will enable more valuable and insightful conclusions to improve the current understand of the link between the sex hormones with asthma, COPD and lung cancer.

A small cohort study of healthy young females demonstrates no change in multiple lung function measurements across all menstrual cycle stages [134]. Although, a minor positive correlation between tidal volume, inspiratory time and expiratory time was reported with estradiol and progesterone during the early-to-mid luteal phase. A study by Hanley in 1981 measured peak expiratory flow rate (PEFR) in 102 female asthmatic patients [135]. Of the 36 patients who reported worsened symptoms at the start of menstruation, PEFR indicated a significant reduction. This indicates that an increase in airway resistance prompted the perception of worse symptoms. However, 65% of the cohort reported no change in symptoms, highlighting that the effect of menstruation on asthma symptoms is inconsistent. A recent similar study combined subjective and objective measurements of premenstrual asthma deterioration in 103 females with asthma. 60% of participants described worsened symptoms in at least one of two menstrual cycles. However, only three females presented with objective deterioration in peak flow rates [136]. An association between the start of the menstrual cycle and asthma symptoms exists; however, there is a discrepancy between the perception of symptoms and physiological changes. Clearly, a highly complex interaction exists between hormone levels, lung physiology and psychological perception of symptoms. Further investigation of this relationship is necessary to improve patient

care and health outcomes and our understanding of disease pathophysiology.

Sex hormones influence the pathophysiology of lung diseases. The exact role and mechanism of how estrogen and testosterone function is yet to be wholly elucidated. The current evidence indicates an association between estrogen and testosterone with clinical symptoms and presentation of these diseases, with no well-defined link to disease development mechanisms between the sexes. There remains a dearth of knowledge surrounding the differential effects of sex hormones in both healthy and disease conditions. Although the role of sex hormones is apparent, a deeper exploration of their signalling and mechanical pathways is required to elucidate how estrogen and testosterone contribute to disease development. The implication of alternate pathogenic factors driving sex differences is evidenced in children where the sex hormonal effect is limited and clear patterns of sex differences exist.

The sex chromosomes

An imbalance exists in disease susceptibility and severity between males and females, which is apparent pre-puberty [6], removing the effects of sex hormones. As such, this draws attention towards fundamental genetic differences between males and females. The concept of sex-biased gene expression is well-established and reviewed in detail by Grath and Parsch [137]. The processes driving sex-biased expression are complex. In particular, sex-chromosome-specific mechanisms such as dosage compensation directly contribute to sexually dimorphic gene expression. Genes on the X and Y chromosome have been shown to contribute to critical cellular processes and are linked to various diseases [138–141]. Therefore, this is a fundamental difference between males and females which may contribute to sex differences in disease susceptibility, progression and severity.

Human cells contain 23 pairs of chromosomes, with 22 pairs referred to as autosomes and the final pair called the sex chromosomes, X and Y. Females have two X-chromosomes (X-chr), and males have one X-chr and one Y-chromosome (Y-chr). The expression of the *SRY* gene from the Y-chr initiates the development of male genitalia, demonstrated by the seminal 'four-core genome' (FCG) mouse model [142]. This model involved transposing *SRY* from the Y-chr to an autosome, meaning that XX and XY mice with ovaries and XX and XY mice with testis could be bred [142]. As a result, it was possible to distinguish whether differences in gene expression from sex chromosome complement drive a sexually dimorphic phenotype or sex hormones [143]. The FCG model has been applied across a range

of experimental designs, which Arnold et al. reviewed in detail and highlighted the importance of this model across different disease systems [143]. In one iteration, distinct differences were observed between XX and XY mice with the same gonadal type, implying a lack of effect by gonadal hormone secretions [144]. Although the effects of estrogen and testosterone must be acknowledged, these models indicate distinct X and Y-chr-specific regulation. In support, studies show that genetic factors drive most differences between the sexes in specific tissues [145, 146].

The sex chromosome complement is a critical biological factor driving sexual dimorphism in disease. The Y-chromosome contains the fewest number of genes (72) compared to the X-chromosome's 833 genes, highlighting a clear imbalance in genotype between males and females. A study investigated the male-specific region (MSY) of the Y-chromosome, identifying unique haplogroups and observed a 50% increased risk for coronary artery disease compared to the other haplotypes [147]. Macrophages from the males carrying the susceptible haplogroup indicated altered processes of inflammation. A wealth of studies and reviews highlight the contribution of the Y-chromosome in protecting against or increasing susceptibility to various diseases [147–150].

The importance of the Y-chr is controversial. It is accepted that the X and Y-chr were once identical, with evolutionary studies demonstrating that they share a common ancestor chromosome. A divergence event causes the Y-chr to undergo significant changes, which results in the partial degeneration of its structure, with some evidence this degradation is continuing [151]. Studies have forecasted that the Y-chromosome is declining, with a steady loss of genes over millions of years. This begged the question, "How important is the Y-chromosome, and will it disappear?" This is a hotly debated topic with two competing schools of thought:

A) Degradation of the Y-chromosome will continue until it eventually becomes extinct
B) Y-chromosome degradation is slowing down, with the remaining genes being critical for survival

The X-chr undergoes a unique process to adjust for this dramatic degradation of the Y-chr, where one X-chr becomes inactivated. X-chr inactivation (XCi) occurs early in human development. Either the maternal or paternal X-chr becomes inactivated in each cell. The same X-chr remains inactivated throughout the mitotic proliferation of that cell [152]. The non-coding RNA "*Xist*" initiates the recruitment of chromatin-modifier proteins, resulting in transcriptional silencing [153].

The silenced X-chr undergoes structural and epigenetic remodelling leading to the formation of a condensed Barr body. This process theoretically accounts for the double dosage of two X chromosomes in females compared to one in males. However, XCi is incomplete, with several genes escaping the inactivation process [146, 154]. Approximately 15–25% of genes escape XCi [155], meaning that these genes are expressed from both X-chromosomes in females. As such, females experience 'double-dosage', whilst males only have a 'single dose' of these genes. Variations in XCi have also been implicated in disease susceptibility [154].

Bellott et al. [156] compared the Y-chromosomes of multiple mammal species to investigate whether an overlap existed for the evolutionarily conserved genes. The authors identified many conserved genes that carried functions beyond sex determination and function to affect all levels of the central dogma—from gene expression to protein translation. Therefore, genes that have survived on the Y-chromosome are critical regulators of various cellular processes. These genes include *RPS4Y1, ZFY, DDX3Y, EIF1AY, KDM5D, KDM6C*. These Y-linked genes have X-chromosome counterparts which are also evolutionarily conserved [156]. However, the sequences of these genes are non-exact, resulting in these proteins having structural variances affecting biological systems through divergent mechanisms and pathways [138, 148,

157]. Therefore, these genes represent fundamental sexual dimorphism at a genetic level, which exists in every cell type. As these genes contribute to the regulation of normal cell and molecular processes at a whole genome level [158], an imbalance of function may contribute to sex differences in disease susceptibility, progression and the response to clinical interventions. The function of some of the sex chromosome-linked homolog genes and their differences are presented in Fig. 5.

Conclusions

Sexual dimorphism is prevalent in various diseases and particularly complex in respiratory diseases—asthma, COPD and lung cancer. These differences affect the susceptibility, severity, presentation and response to medical treatments. Despite the increased study of the factors contributing to sex differences in recent years, more research is required. The sex hormones estrogen and testosterone are well-recognised to contribute to the severity of disease, but how and whether they are the primary factors causing disease pathogenesis remains unclear. We have described the phenomena of dosage compensation and XCi escape causing an imbalance of key genome regulators from the sex chromosomes. These genes contribute to central disease features such as inflammation and fibrosis. Therefore, they are valuable candidates to further our understanding of
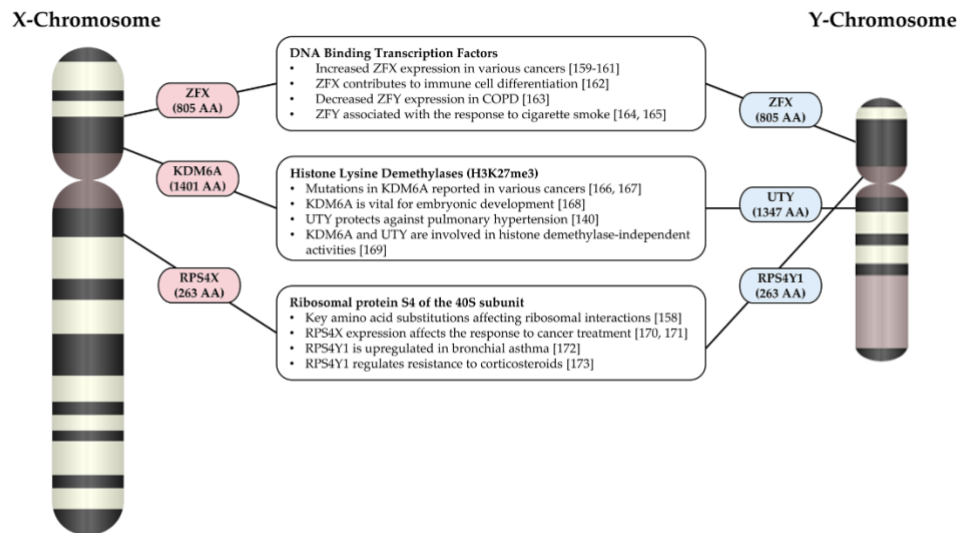
Fig. 5 Illustration of genome regulators on the X and Y chromosome and their contribution to various diseases. *AA* amino acids, *H3K27me3* Histone 3 Lysine residue 27 trimethylation [140, 158–171, 171–173]

the development of disease and the generation of new clinical interventions to improve the health outcomes for males and females.

Abbreviations

COPD	chronic obstructive pulmonary disease
FEV_1	forced expiratory volume in one second
ICS	inhaled corticosteroid
TNFα	tumour necrosis factor alpha
ECM	extracellular matrix
FAO	fixed airflow obstruction
AHR	airway hyperresponsiveness
TGFβ	transformation growth factor beta
SEO-COPD	Severe early onset chronic obstructive pulmonary disease
NSCLC	non-small cell lung cancer
SCLC	small cell lung cancer
ERT	estrogen replacement therapy
HRT	hormone replacement therapy
PEFR	peak expiratory flow rate
X-chr	X chromosome
Y-chr	Y chromosome
FCG	four core genome
XCi	X chromosome inactivation

Availability of data and materials
Not applicable.

Declarations

Ethics approval and consent to participate
Protocols were submitted to and approved by a human research ethics committee and prior written and informed consent was obtained from patients under approval by code #X14-0045.

Consent for publication
Not applicable.

Competing interests
The authors declare that there are no competing of interests, financial or otherwise.

References
1. Deegan DF, Engel N. Sexual dimorphism in the age of genomics: how, when, where. Front Cell Dev Biol. 2019;7:186.
2. Glasgow AM, Greene CM. Epigenetic mechanisms underpinning sexual dimorphism in lung disease. Future Med. 2022;14:65–7.
3. Rinn JL, Snyder M. Sexual dimorphism in mammalian gene expression. Trends Genet. 2005;21(5):298–305.
4. National Institute of Health. NOT-OD-15-102: consideration of sex as a biological variable in NIH-funded research. 2015.
5. Clayton JA, Tannenbaum C. Reporting sex, gender, or both in clinical research? JAMA. 2016;316(18):1863–4.
6. Reddy KD, Oliver BG. Sex-specific effects of in utero and adult tobacco smoke exposure. Am J Physiol Lung Cell Mol Physiol. 2021;320(1):L63–72.
7. Thompson K, Venkatesh B, Hammond N, Taylor C, Finfer S. Sex differences in response to adjunctive corticosteroid treatment for patients with septic shock. Intensive Care Med. 2021;47(2):246–8.
8. Mersha TB, Martin LJ, Myers JMB, Kovacic MB, He H, Lindsey M, Sivaprasad U, Chen W, Hershey GKK. Genomic architecture of asthma differs by sex. Genomics. 2015;106(1):15–22.
9. Gautam Y, Afanador Y, Abebe T, López JE, Mersha TB. Genome-wide analysis revealed sex-specific gene expression in asthmatics. Hum Mol Genet. 2019;28(15):2600–14.
10. Jenkins CR, Boulet L-P, Lavoie KL, Raherison-Semjen C, Singh D. Personalised treatment of asthma: the importance of sex and gender differences. J Allergy Clin Immunol. 2022. https://doi.org/10.1016/j.jaip.2022.02.002.
11. Yu XQ, Yap ML, Cheng ES, Ngo PJ, Vaneckova P, Karikios D, Canfell K, Weber MF. Evaluating prognostic factors for sex differences in lung cancer survival: findings from a large Australian cohort. J Thorac Oncol. 2022;17(5):688–99.
12. Bazan IS, Kim S-J, Ardito TA, Zhang Y, Shan P, Sauler M, Lee PJ. Sex differences and altered mitophagy in experimental pulmonary hypertension. Am J Physiol Lung Cell Mol Physiol. 2022. https://doi.org/10.1152/ajplung.00019.2020.
13. LoMauro A, Aliverti A. Sex differences in respiratory function. Breathe. 2018;14(2):131–40.
14. Prakash Y, Pabelick CM, Chiarella SE. Sex differences in respiratory physiology. Sex-based differences in lung physiology. 2021: 1–11 https://doi.org/10.1007/978-3-030-63549-7_1
15. Han S, Mallampalli RK. The role of surfactant in lung disease and host defense against pulmonary infections. Ann Am Thorac Soc. 2015;12(5):765–74.
16. Farrell PM, Avery ME. Hyaline membrane disease. Am Rev Respir Dis. 1975;111(5):657–88.
17. Seaborn T, Simard M, Provost PR, Piedboeuf B, Tremblay Y. Sex hormone metabolism in lung development and maturation. Trends Endocrinol Metab. 2010;21(12):729–38.
18. Sheel AW, Dominelli PB, Molgat-Seon Y. Revisiting dysanapsis: sex-based differences in airways and the mechanics of breathing during exercise. Exp Physiol. 2016;101(2):213–8.
19. Green M, Mead J, Turner JM. Variability of maximum expiratory flow-volume curves. J Appl Physiol. 1974;37(1):67–74.
20. Thompson BR. Dysanapsis—once believed to be a physiological curiosity—is now clinically important. Am Thorac Soc. 2017;195:277–8.
21. Hibbert M, Lannigan A, Raven J, Landau L, Phelan P. Gender differences in lung growth. Pediatr Pulmonol. 1995;19(2):129–34.
22. Newcomb DC, Shah R. Sex bias in asthma prevalence and pathogenesis. Front Immunol. 2018;9:2997.
23. Ekpruke CD, Silveyra P. Sex differences in airway remodeling and inflammation: clinical and biological factors. Front Allergy. 2022. https://doi.org/10.3389/falgy.2022.875295.
24. Han M, Murray S, Fell CD, Flaherty KR, Toews GB, Myers J, Colby TV, Travis WD, Kazerooni EA, Gross BH. Sex differences in physiological progression of idiopathic pulmonary fibrosis. Eur Respir J. 2008;31(6):1183–8.
25. Lam GY, Goodwin J, Wilcox PG, Quon BS. Sex disparities in cystic fibrosis: review on the effect of female sex hormones on lung pathophysiology and outcomes. ERJ Open Res. 2021;7:1.
26. Asthma Gif. Global Strategy for Asthma Managemnet and Prevention. Global Initiative for Asthma. 2022.
27. Rasmussen F, Taylor DR, Flannery EM, Cowan JO, Greene JM, Herbison GP, Sears MR. Risk factors for airway remodeling in asthma manifested by a low postbronchodilator FEV1/vital capacity ratio: a longitudinal

13

population study from childhood to adulthood. Am J Respir Crit Care Med. 2002;165(11):1480–8.

28. Chowdhury NU, Guntur VP, Newcomb DC, Wechsler ME. Sex and gender in asthma. Eur Respir Rev. 2021;30:162.

29. Australian Bureau of Statistics (2017–18), National Health Survey: First Results, ABS website: [https://www.abs.gov.au/statistics/health/health-conditions-and-risks/national-health-survey-first-results/2017-18], Accessed 14 December 2022.

30. Australian Institute of Health and Welfare. 2022, Australia's children, AIHW, Asutralian Government, [https://www.aihw.gov.au/reports/chronic-respiratory-conditions/asthma/data], Accessed 14 December 2022.

31. Grainge CL, Maltby S, Gibson PG, Wark PA, McDonald VM. Targeted therapeutics for severe refractory asthma: monoclonal antibodies. Expert Rev Clin Pharmacol. 2016;9(7):927–41.

32. Fuseini H, Newcomb DC. Mechanisms driving gender differences in asthma. Curr Allergy Asthma Rep. 2017;17(3):1–9.

33. Raherison C, Janson C, Jarvis D, Burney P, Cazzoletti L, de Marco R, Neukirch F, Leynaert B. Evolution of asthma severity in a cohort of young adults: is there any gender difference? PLoS ONE. 2009;4(9):e7146.

34. Ricciardolo FLM, Levra S, Sprio AE, Bertolini F, Carriero V, Gallo F, Ciprandi G. Asthma in the real-world: the relevance of gender. Int Arch Allergy Immunol. 2020;181(6):462–6.

35. Ambhore NS, Kalidhindi RSR, Sathish V. Sex-steroid signaling in lung diseases and inflammation. Lung Inflamm Health Dis. 2021;1303:243.

36. Millas I, Duarte Barros M. Estrogen receptors and their roles in the immune and respiratory systems. Anat Rec. 2021;304(6):1185–93.

37. Aravamudan B, Goorhouse KJ, Unnikrishnan G, Thompson MA, Pabelick CM, Hawse JR, Prakash Y, Sathish V. Differential expression of estrogen receptor variants in response to inflammation signals in human airway smooth muscle. J Cell Physiol. 2017;232(7):1754–60.

38. Espuela-Ortiz A, Herrera-Luis E, Lorenzo-Díaz F, Hu D, Eng C, Villar J, Rodriguez-Santana JR, Burchard EG, Pino-Yanes M. Role of sex on the genetic susceptibility to childhood asthma in latinos and African Americans. J Pers Med. 2021;11(11):1140.

39. Osman M. Therapeutic implications of sex differences in asthma and atopy. Arch Dis Child. 2003;88(7):587–90.

40. Genuneit J. Sex-specific development of asthma differs between farm and nonfarm children: a cohort study. Am J Respir Crit Care Med. 2014;190(5):588–90.

41. Naeem A, Silveyra P. Sex differences in paediatric and adult asthma. Eur Med J. 2019;4(2):27.

42. Bijanzadeh M, Mahesh PA, Ramachandra NB. An understanding of the genetic basis of asthma. Indian J Med Res. 2011;134(2):149.

43. Chiang C-H, Chuang C-H, Liu S-L, Shen H-D. Genetic polymorphism of transforming growth factor β1 and tumor necrosis factor α is associated with asthma and modulates the severity of asthma. Respir Care. 2013;58(8):1343–50.

44. Malerba G, Pignatti PF. A review of asthma genetics: gene expression studies and recent candidates. J Appl Genet. 2005;46(1):93–104.

45. Shin HD, Park BL, Kim LH, Jung JH, Wang HJ, Kim YJ, Park H-S, Hong S-J, Choi BW, Kim D-J. Association of tumor necrosis factor polymorphisms with asthma and serum total IgE. Hum Mol Genet. 2004;13(4):397–403.

46. Uppal S, Verma S, Dhot P. Normal values of CD4 and CD8 lymphocyte subsets in healthy Indian adults and the effects of sex, age, ethnicity, and smoking. Cytometry B Clin Cytom. 2003;52(1):32–6.

47. Los H, Koppelman G, Postma D. The importance of genetic influences in asthma. Eur Respir J. 1999;14(5):1210–27.

48. Gon Y, Hashimoto S. Role of airway epithelial barrier dysfunction in pathogenesis of asthma. Allergol Int. 2018;67(1):12–7.

49. Barczyk A, Pierzchala W, Sozanska E. Interleukin-17 in sputum correlates with airway hyperresponsiveness to methacholine. Respir Med. 2003;97(6):726–33.

50. Miyasaka T, Dobashi-Okuyama K, Kawakami K, Masuda-Suzuki C, Takayanagi M, Ohno I. Sex plays a multifaceted role in asthma pathogenesis. Biomolecules. 2022;12(5):650.

51. Jevnikar Z, Östling J, Ax E, Calvén J, Thörn K, Israelsson E, Öberg L, Singhania A, Lau LC, Wilson SJ. Epithelial IL-6 trans-signaling defines a new asthma phenotype with increased airway inflammation. J Allergy Clin Immunol. 2019;143(2):577–90.

52. King GG, James A, Harkness L, Wark PA. Pathophysiology of severe asthma: we've only just started. Respirology. 2018;23(3):262–71.

53. Araujo BB, Dolhnikoff M, Silva LF, Elliot J, Lindeman J, Ferreira D, Mulder A, Gomes HA, Fernezlian S, James A. Extracellular matrix components and regulators in the airway smooth muscle in asthma. Eur Respir J. 2008;32(1):61–9.

54. Ito JT, Lourenço JD, Righetti RF, Tibério IF, Prado CM, Lopes FD. Extracellular matrix component remodeling in respiratory diseases: what has been found in clinical and experimental studies? Cells. 2019;8(4):342.

55. Burgess JK, Boustany S, Moir LM, Weckmann M, Lau JY, Grafton K, Baraket M, Hansbro PM, Hansbro NG, Foster PS. Reduction of tumstatin in asthmatic airways contributes to angiogenesis, inflammation, and hyperresponsiveness. Am J Respir Crit Care Med. 2010;181(2):106–15.

56. Zhang L, He L, Gong J, Liu C. Risk factors associated with irreversible airway obstruction in asthma: a systematic review and meta-analysis. BioMed Res Int. 2016. https://doi.org/10.1155/2016/9868704.

57. Takeda M, Tanabe M, Ito W, Ueki S, Konnno Y, Chihara M, Itoga M, Kobayashi Y, Moritoki Y, Kayaba H. Gender difference in allergic airway remodelling and immunoglobulin production in mouse model of asthma. Respirology. 2013;18(5):797–806.

58. (GOLD) GlfCOLD. Pocket guide to COPD diagnosis management and prevention global intiative for chronic obstructive lung disease. New Delhi: INC; 2019.

59. Miravitlles M, Ribera A. Understanding the impact of symptoms on the burden of COPD. Respir Res. 2017;18(1):1–11.

60. Australian Institute of Health and Welfare (2023) Chronic respiratory conditions, AIHW, Australian Government [https://www.aihw.gov.au/reports/chronic-respiratory-conditions/copd/contents/copd], Accessed 10 December 2022.

61. Akinbami OJ, Liu X. Chronic obstuctive pulmonary disease among adults aged 18 and over in the United States, 1998–2009. Washington DC: US Department of Health and Human Services Centers for Disease Control and; 2011.

62. Barnes PJ. Sex differences in chronic obstructive pulmonary disease mechanisms. In.: American Thoracic Society; 2016.

63. Tam A, Churg A, Wright JL, Zhou S, Kirby M, Coxson HO, Lam S, Man SP, Sin DD. Sex differences in airway remodeling in a mouse model of chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2016;193(8):825–34.

64. Gan WQ, Man SP, Postma DS, Camp P, Sin DD. Female smokers beyond the perimenopausal period are at increased risk of chronic obstructive pulmonary disease: a systematic review and meta-analysis. Respir Res. 2006;7(1):1–9.

65. Eisner MD, Balmes J, Katz PP, Trupin L, Yelin EH, Blanc PD. Lifetime environmental tobacco smoke exposure and the risk of chronic obstructive pulmonary disease. Environ Health. 2005;4(1):7.

66. Sørheim I-C, Johannessen A, Gulsvik A, Bakke PS, Silverman EK, DeMeo DL. Gender differences in COPD: are women more susceptible to smoking effects than men? Thorax. 2010;65(6):480–5.

67. Foreman MG, Zhang L, Murphy J, Hansel NN, Make B, Hokanson JE, Washko G, Regan EA, Crapo JD, Silverman EK. Early-onset chronic obstructive pulmonary disease is associated with female sex, maternal factors, and African American race in the COPDGene study. Am J Respir Crit Care Med. 2011;184(4):414–20.

68. Fazleen A, Wilkinson T. Early COPD: current evidence for diagnosis and management. Ther Adv Respir Dis. 2020;14:1753466620942128.

69. Kohler M, Sandberg A, Kjellqvist S, Thomas A, Karimi R, Nyrén S, Eklund A, Thevis M, Sköld CM, Wheelock ÅM. Gender differences in the bronchoalveolar lavage cell proteome of patients with chronic obstructive pulmonary disease. J Allergy Clin Immunol. 2013;131(3):743–51.

70. Haspel JA, Choi AM. Autophagy: a core cellular process with emerging links to pulmonary disease. Am J Respir Crit Care Med. 2011;184(11):1237–46.

71. Daffa N, Tighe PJ, Corne J, Fairclough LC, Todd I. Natural and disease-specific autoantibodies in chronic obstructive pulmonary disease. Clin Exp Immunol. 2015;180(1):155–63.

72. Whitacre CC. Sex differences in autoimmune disease. Nat Immunol. 2001;2(9):777–80.

73. Forsslund H, Yang M, Mikko M, Karimi R, Nyrén S, Engvall B, Grunewald J, Merikallio H, Kaarteenaho R, Wahlström J. Gender differences in the T-cell profiles of the airways in COPD patients associated with clinical phenotypes. Int J Chron Obstruct Pulmon Dis. 2016. https://doi.org/10.2147/COPD.S113625.

74. Pirker R. Conquering lung cancer: current status and prospects for the future. Pulmonology. 2020;26(5):283–90.
75. Cruz CSD, Tanoue LT, Matthay RA. Lung cancer: epidemiology, etiology, and prevention. Clin Chest Med. 2011;32(4):605–44.
76. Ragavan M, Patel MI. The evolving landscape of sex-based differences in lung cancer: a distinct disease in women. Eur Respir Rev. 2022;31:163.
77. Stabellini N, Bruno DS, Dmukauskas M, Barda AJ, Cao L, Shanahan J, Waite K, Montero AJ, Barnholtz-Sloan JS. Sex differences in lung cancer treatment and outcomes at a large hybrid academic-community practice. JTO Clin Res Rep. 2022;3(4):100307.
78. Stapelfeld C, Dammann C, Maser E. Sex-specificity in lung cancer risk. Int J Cancer. 2020;146(9):2376–82.
79. Karachaliou N, Pilotto S, Lazzari C, Bria E, de Marinis F, Rosell R. Cellular and molecular biology of small cell lung cancer: an overview. Transl Lung Cancer Res. 2016;5(1):2.
80. Cooper WA, Lam DC, O'Toole SA, Minna JD. Molecular biology of lung cancer. J Thorac Dis. 2013;5(Suppl 5):S479.
81. Larsen JE, Minna JD. Molecular biology of lung cancer: clinical implications. Clin Chest Med. 2011;32(4):703–40.
82. Zang EA, Wynder EL. Differences in lung cancer risk between men and women: examination of the evidence. JNCI J Natl Cancer Inst. 1996;88(3–4):183–92.
83. Ragavan MV, Patel MI. Understanding sex disparities in lung cancer incidence: are women more at risk? Future Med. 2020. https://doi.org/10.2217/lmt-2020-0013.
84. Thun MJ, Hannan LM, Adams-Campbell LL, Boffetta P, Buring JE, Feskanich D, Flanders WD, Jee SH, Katanoda K, Kolonel LN. Lung cancer occurrence in never-smokers: an analysis of 13 cohorts and 22 cancer registry studies. PLoS Med. 2008;5(9):e185.
85. Sagerup CM, Småstuen M, Johannesen TB, Helland Å, Brustugun OT. Sex-specific trends in lung cancer incidence and survival: a population study of 40 118 cases. Thorax. 2011;66(4):301–7.
86. Fidler-Benaoudia MM, Torre LA, Bray F, Ferlay J, Jemal A. Lung cancer incidence in young women vs. young men a systematic analysis in 40 countries. Int J Cancer. 2020;147(3):811–9.
87. Tong BC, Kosinski AS, Burfeind WR Jr, Onaitis MW, Berry MF, Harpole DH Jr, D'Amico TA. Sex differences in early outcomes after lung cancer resection: analysis of the society of thoracic surgeons general thoracic database. J Thorac Cardiovasc Surg. 2014;148(1):13–8.
88. Pinto JA, Vallejos CS, Raez LE, Mas LA, Ruiz R, Torres-Roman JS, Morante Z, Araujo JM, Gómez HL, Aguilar A. Gender and outcomes in non-small cell lung cancer: an old prognostic variable comes back for targeted therapy and immunotherapy? ESMO Open. 2018;3(3): e000344.
89. Liang J, Hong J, Tang X, Qiu X, Zhu K, Zhou L, Guo D. Sex difference in response to non-small cell lung cancer immunotherapy: an updated meta-analysis. Ann Med. 2022;54(1):2606–16.
90. Australian Institute of Health and Welfare (2022) **Cancer data in Australia**, AIHW, Australian Government, [https://www.aihw.gov.au/reports/cancer/cancer-data-in-australia/data], Accessed 16 December 2022.
91. Cook MB, McGlynn KA, Devesa SS, Freedman ND, Anderson WF. Sex disparities in cancer mortality and survivalsexs disparities in cancer mortality. Cancer Epidemiol Biomark Prev. 2011;20(8):1629–37.
92. Gu Y, Tang YY, Wan JX, Zou JY, Lu CG, Zhu HS, Sheng SY, Wang YF, Liu HC, Yang J. Sex difference in the expression of PD-1 of non-small cell lung cancer. Front Immunol. 2022. https://doi.org/10.3389/fimmu.2022.1026214.
93. Wierman ME. Sex steroid effects at target tissues: mechanisms of action. Adv Physiol Educ. 2007;31(1):26–33.
94. Tam A, Morrish D, Wadsworth S, Dorscheid D, Man SP, Sin DD. The role of female hormones on lung function in chronic lung diseases. BMC Womens Health. 2011;11(1):24.
95. LoMauro A, Aliverti A. Sex and gender in respiratory physiology. Eur Respir Rev. 2021;30:162.
96. Bulkhi AA, Shepard KV II, Casale TB, Cardet JC. Elevated testosterone is associated with decreased likelihood of current asthma regardless of sex. J Allergy Clin Immunol. 2020;8(9):3029–35.
97. Zein JG, McManus JM, Sharifi N, Erzurum SC, Marozkina N, Lahm T, Giddings O, Davis MD, DeBoer MD, Comhair SA. Benefits of airway androgen receptor expression in human asthma. Am J Respir Crit Care Med. 2021;204(3):285–93.
98. Montano LM, Espinoza J, Flores-Soto E, Chávez J, Perusquia M. Androgens are bronchoactive drugs that act by relaxing airway smooth muscle and preventing bronchospasm. J Endocrinol. 2014;222(1):1–13.
99. Dijkstra A, Howard TD, Vonk JM, Ampleford EJ, Lange LA, Bleecker ER, Meyers DA, Postma DS. Estrogen receptor 1 polymorphisms are associated with airway hyperresponsiveness and lung function decline, particularly in female subjects with asthma. J Allergy Clin Immunol. 2006;117(3):604–11.
100. Carey MA, Card JW, Bradbury JA, Moorman MP, Haykal-Coates N, Gavett SH, Graves JP, Walker VR, Flake GP, Voltz JW. Spontaneous airway hyper-responsiveness in estrogen receptor-α–deficient mice. Am J Respir Crit Care Med. 2007;175(2):126–35.
101. Radzikowska U, Golebski K. Sex hormones and asthma: the role of estrogen in asthma development and severity. Hoboken: Wiley Online Library; 2022.
102. Holguin F. Sex hormones and asthma. Am Thorac Soc. 2020;201:127–8.
103. Rao CK, Moore CG, Bleecker E, Busse WW, Calhoun W, Castro M, Chung KF, Erzurum SC, Israel E, Curran-Everett D. Characteristics of perimenstrual asthma and its relation to asthma severity and control: data from the severe asthma research program. Chest. 2013;143(4):984–92.
104. Macsali F, Svanes C, Sothern RB, Benediktsdottir B, Bjørge L, Dratva J, Franklin KA, Holm M, Janson C, Johannessen A. Menstrual cycle and respiratory symptoms in a general nordic-baltic population. Am J Respir Crit Care Med. 2013;187(4):366–73.
105. Balzano G, Fuschillo S, Melillo G, Bonini S. Asthma and sex hormones. Allergy. 2001;56(1):13–20.
106. Raghavan D, Jain R. Increasing awareness of sex differences in airway diseases. Respirology. 2016;21(3):449–59.
107. Chiarella SE, Cardet JC, Prakash Y. Sex, cells, and asthma. Mayo Clin Proc. 2021. https://doi.org/10.1016/j.mayocp.2020.12.007.
108. Fuentes N, Silva Rodriguez M, Silveyra P. Role of sex hormones in lung cancer. Exp Biol Med. 2021. https://doi.org/10.1177/15353702211019697.
109. Mikkonen L, Pihlajamaa P, Sahu B, Zhang F-P, Jänne OA. Androgen receptor and androgen-dependent gene expression in lung. Mol Cell Endocrinol. 2010;317(1–2):14–24.
110. Costa AR, de Oliveira ML, Cruz I, Gonçalves I, Cascalheira JF, Santos CR. The sex bias of cancer. Trends Endocrinol Metab. 2020. https://doi.org/10.1016/j.tem.2020.07.002.
111. Townsend EA, Thompson MA, Pabelick CM, Prakash Y. Rapid effects of estrogen on intracellular Ca2+ regulation in human airway smooth muscle. Am J Physiol Lung Cell Mol Physiol. 2010;298(4):L521–30.
112. Han Y-Y, Forno E, Celedón JC. Sex steroid hormones and asthma in a nationwide study of US adults. Am J Respir Crit Care Med. 2019. https://doi.org/10.1164/rccm.201905-0996OC.
113. Kwon HL, Belanger K, Bracken MB. Asthma prevalence among pregnant and childbearing-aged women in the United States: estimates from national health surveys. Ann Epidemiol. 2003;13(5):317–24.
114. Sawicki E, Stewart K, Wong S, Paul E, Leung L, George J. Management of asthma by pregnant women attending an Australian maternity hospital. Aust N Z J Obstet Gynaecol. 2012;52(2):183–8.
115. Murphy VE, Clifton VL, Gibson PG. Asthma exacerbations during pregnancy: incidence and association with adverse pregnancy outcomes. Thorax. 2006;61(2):169–76.
116. Wang H, Li N, Huang H. Asthma in pregnancy: pathophysiology, diagnosis, whole-course management, and medication safety. Can Respir J. 2020. https://doi.org/10.1155/2020/9046842.
117. Schatz M, Harden K, Forsythe A, Chilingar L, Hoffman C, Sperling W, Zeiger RS. The course of asthma during pregnancy, post partum, and with successive pregnancies: a prospective analysis. J Allergy Clin Immunol. 1988;81(3):509–17.
118. Gothi D, Sah RB, Teotia A, Yadav S. Improvement in spirometry and oxygenation of chronic obstructive pulmonary disease during pregnancy. Lung India. 2018;35(5):441.
119. Lalli CM, Raju L. Pregnancy and chronic obstructive pulmonary disease. Chest. 1981;80(6):759–61.
120. Soares A, Dos Santos J, Silva A, Magalhães H, Estevinho F, Sottomayor C. Treatment of lung cancer during pregnancy. Pulmonology. 2020;26(5):314–7.
121. Hsu L-H, Chu N-M, Kao S-H. Estrogen, estrogen receptor and lung cancer. Int J Mol Sci. 2017;18(8):1713.

15

122. Stensheim H, Møller B, Van Dijk T, Fosså SD. Cause-specific survival for women diagnosed with cancer during pregnancy or lactation: a registry-based cohort study. J Clin Oncol. 2009;27(1):45–51.

123. Campbell B, Davis S, Abramson M, Mishra G, Handelsman D, Perret J, Dharmage S. Menopause, lung function and obstructive lung disease outcomes: a systematic review. Climacteric. 2018;21(1):3–12.

124. Troisi RJ, Speizer FE, Willett WC, Trichopoulos D, Rosner B. Menopause, postmenopausal estrogen preparations, and the risk of adult-onset asthma a prospective cohort study. Am J Respir Crit Care Med. 1995;152(4):1183–8.

125. Scioscia G, Carpagnano GE, Lacedonia D, Soccio P, Quarato CMI, Trabace L, Fuso P, Foschino Barbaro MP. The role of airways 17β-estradiol as a biomarker of severity in postmenopausal asthma: a pilot study. J Clin Med. 2020;9(7):2037.

126. Kaufman J-M, Lapauw B, Mahmoud A, T'Sjoen G, Huhtaniemi IT. Aging and the male reproductive system. Endocr Rev. 2019;40(4):906–72.

127. Real FG, Svanes C, Omenaas ER, Antò JM, Plana E, Jarvis D, Janson C, Neukirch F, Zemp E, Dratva J. Lung function, respiratory symptoms, and the menopausal transition. J Allergy Clin Immunol. 2008;121(1):72–80.

128. Songür N, Aydin ZD, Oztürk Ö, Sahin Ü, Khayri U, Bircan A, Akkaya A. Respiratory symptoms, pulmonary function, and reproductive history: isparta menopause and health study. J Womens Health. 2010;19(6):1145–54.

129. Lange P, Parner J, Prescott E, Ulrik CS, Vestbo J. Exogenous female sex steroid hormones and risk of asthma and asthma-like symptoms: a cross sectional study of the general population. Thorax. 2001;56(8):613–6.

130. Chung H-F, Gete DG, Mishra GD. Age at menopause and risk of lung cancer: a systematic review and meta-analysis. Maturitas. 2021;153:1–10.

131. Meinhold CL, de Berrington González A, Bowman ED, Brenner AV, Jones RT, Lacey JV Jr, Loffredo CA, Perlmutter D, Schonfeld SJ, Trivers GE. Reproductive and hormonal factors and the risk of nonsmall cell lung cancer. Int J Cancer. 2011;128(6):1404–13.

132. Jin K, Hung RJ, Thomas S, Le Marchand L, Matsuo K, Seow A, Shen H, Kok WP, Yuan JM, Wu M. Hormonal factors in association with lung cancer among Asian women: a pooled analysis from the international lung cancer consortium. Int J Cancer. 2021;148(9):2241–54.

133. Miller K. Estrogen and DNA damage: the silent source of breast cancer? J Natl Cancer Inst. 2003;95(2):100–2.

134. da Silva SB, Viana EdSR, de Sousa MBC. Changes in peak expiratory flow and respiratory strength during the menstrual cycle. Respir Physiol Neurobiol. 2006;150(2–3):211–9.

135. Hanley S. Asthma variation with menstruation. Br J Dis Chest. 1981;75(3):306–8.

136. Vega AP, Ramos JS, Pérez JM, Gutierrez FA, García JI, Oliva RV, Palacios PR, Nieto JB, Rodríguez IS, Muñoz FG. Variability in the prevalence of premenstrual asthma. Eur Respir J. 2010;35(5):980–6.

137. Grath S, Parsch J. Sex-biased gene expression. Annu Rev Genet. 2016;50:29–44.

138. Lau Y-FC. Y chromosome in health and diseases. Cell Biosci. 2020;10(1):1–10.

139. Du S, Itoh N, Askarinam S, Hill H, Arnold AP, Voskuhl RR. XY sex chromosome complement, compared with XX, in the CNS confers greater neurodegeneration during experimental autoimmune encephalomyelitis. Proc Natl Acad Sci. 2014;111(7):2806–11.

140. Cunningham CM, Li M, Ruffenach G, Doshi M, Aryan L, Hong J, Park J, Hrncir H, Medzikovic L, Umar S. Y-chromosome gene, uty, protects against pulmonary hypertension by reducing proinflammatory chemokines. Am J Respir Critical Care Med. 2022. https://doi.org/10.1164/rccm.202110-2309OC.

141. Link JC, Wiese CB, Chen X, Avetisyan R, Ronquillo E, Ma F, Guo X, Yao J, Allison M, Chen Y-DI. X chromosome dosage of histone demethylase KDM5C determines sex differences in adiposity. J Clin Invest. 2020;130:11.

142. De Vries GJ, Rissman EF, Simerly RB, Yang L-Y, Scordalakes EM, Auger CJ, Swain A, Lovell-Badge R, Burgoyne PS, Arnold AP. A model system for study of sex chromosome effects on sexually dimorphic neural and behavioral traits. J Neurosci. 2002;22(20):9005–14.

143. Arnold AP, Chen X. What does the "four core genotypes" mouse model tell us about sex differences in the brain and other tissues? Front Neuroendocrinol. 2009;30(1):1–9.

144. Gioiosa L, Chen X, Watkins R, Klanfer N, Bryant CD, Evans CJ, Arnold AP. Sex chromosome complement affects nociception in tests of acute and chronic exposure to morphine in mice. Horm Behav. 2008;53(1):124–30.

145. Ober C, Loisel DA, Gilad Y. Sex-specific genetic architecture of human disease. Nat Rev Genet. 2008;9(12):911.

146. Arnold AP, Reue K, Eghbali M, Vilain E, Chen X, Ghahramani N, Itoh Y, Li J, Link JC, Ngun T. The importance of having two X chromosomes. Philos Trans R Soc B Biol Sci. 2016;371(1688):20150113.

147. Charchar FJ, Bloomer LD, Barnes TA, Cowley MJ, Nelson CP, Wang Y, Denniff M, Debiec R, Christofidou P, Nankervis S. Inheritance of coronary artery disease in men: an analysis of the role of the Y chromosome. Lancet. 2012;379(9819):915–22.

148. Wilson MA. The Y chromosome and its impact on health and disease. Hum Mol Genet. 2021;30(R2):R296–300.

149. Umar S, Cunningham CM, Itoh Y, Moazeni S, Vaillancourt M, Sarji S, Centala A, Arnold AP, Eghbali M. The Y chromosome plays a protective role in experimental hypoxic pulmonary hypertension. Am J Respir Crit Care Med. 2018;197(7):952–5.

150. Arnold AP. Y chromosome's roles in sex differences in disease. Proc Natl Acad Sci US A. 2017;114(15):3787–9.

151. Griffin DK. Is the Y chromosome disappearing?—both sides of the argument. Chromosome Res. 2012;20(1):35–45.

152. Panning B. X-chromosome inactivation: the molecular basis of silencing. J Biol. 2008;7(8):1–4.

153. Chlamydas S, Markouli M, Strepkos D, Piperi C. Epigenetic mechanisms regulate sex-specific bias in disease manifestations. J Mol Med. 2022. https://doi.org/10.1007/s00109-022-02227-x.

154. Shvetsova E, Sofronova A, Monajemi R, Gagalova K, Draisma HH, White SJ, Santen GW, de Sousa Chuva, Lopes SM, Heijmans BT, van Meurs J. Skewed X-inactivation is common in the general female population. Eur J Human Genet. 2019;27(3):455–65.

155. Balaton BP, Cotton AM, Brown CJ. Derivation of consensus inactivation status for X-linked genes from genome-wide studies. Biol Sex Differ. 2015;6(1):1–11.

156. Bellott DW, Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Cho T-J, Koutseva N, Zaghlul S, Graves T, Rock S. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. Nature. 2014;508(7497):494.

157. Meester I, Manilla-Muñoz E, León-Cachón RB, Paniagua-Frausto GA, Carrión-Alvarez D, Ruiz-Rodríguez CO, Rodríguez-Rangel X, García-Martínez JM. SeXY chromosomes and the immune system: reflections after a comparative study. Biol Sex Differ. 2020;11(1):1–13.

158. Lopes AM, Miguel RN, Sargent CA, Ellis PJ, Amorim A, Affara NA. The human RPS4 paralogue on Yq11. 223 encodes a structurally conserved ribosomal protein and is preferentially expressed during spermatogenesis. BMC Mol Biol. 2010;11(1):1–12.

159. Weisberg SP, Smith-Raska MR, Esquilin JM, Zhang J, Arenzana TL, Lau CM, Churchill M, Pan H, Klinakis A, Dixon JE. ZFX controls propagation and prevents differentiation of acute T-lymphoblastic and myeloid leukemia. Cell Rep. 2014;6(3):528–40.

160. Rhie SK, Yao L, Luo Z, Witt H, Schreiner S, Guo Y, Perez AA, Farnham PJ. ZFX acts as a transcriptional activator in multiple types of human tumors by binding downstream from transcription start sites at the majority of CpG island promoters. Genome Res. 2018;28(3):310–20.

161. Zhu Z, Li K, Xu D, Liu Y, Tang H, Xie Q, Xie L, Liu J, Wang H, Gong Y. ZFX regulates glioma cell proliferation and survival in vitro and in vivo. J Neurooncol. 2013;112(1):17–25.

162. Galan-Caridad JM, Harel S, Arenzana TL, Hou ZE, Doetsch FK, Mirny LA, Reizis B. Zfx controls the self-renewal of embryonic and hematopoietic stem cells. Cell. 2007;129(2):345–57.

163. van den Berge M, Brandsma C-A, Faiz A, de Vries M, Rathnayake SN, Paré PD, Sin DD, Bossé Y, Laviolette M, Nickle DC. Differential lung tissue gene expression in males and females: implications for the susceptibility to develop COPD. Eur Respir J. 2019. https://doi.org/10.1183/13993003.02567-2017.

164. Banday AR, Papenberg BW, Prokunina-Olsson L. When the smoke clears m6A from a Y chromosome-linked lncRNA, men get an increased risk of cancer. Can Res. 2020;80(13):2718–9.

165. Irimie AI, Braicu C, Cojocneanu R, Magdo L, Onaciu A, Ciocan C, Mehterov N, Dudea D, Buduru S, Berindan-Neagoe I. Differential effect of smoking on gene expression in head and neck cancer patients. Int J Environ Res Public Health. 2018;15(7):1558.

166. Gažová I, Lengeling A, Summers KM. Lysine demethylases KDM6A and UTY: The X and Y of histone demethylation. Mol Genet Metab. 2019;127(1):31–44.

167. Tricarico R, Nicolas E, Hall MJ, Golemis EA. X-and Y-linked chromatin-modifying genes as regulators of sex-specific cancer incidence and prognosis. Clin Cancer Res. 2020. https://doi.org/10.1158/1078-0432. CCR-20-1741.

168. Welstead GG, Creyghton MP, Bilodeau S, Cheng AW, Markoulaki S, Young RA, Jaenisch R. X-linked H3K27me3 demethylase Utx is required for embryonic development in a sex-specific manner. Proc Natl Acad Sci. 2012;109(32):13004–9.

169. Shpargel KB, Sengoku T, Yokoyama S, Magnuson T. UTX and UTY demonstrate histone demethylase-independent function in mouse embryonic development. PLoS Genet. 2012. https://doi.org/10.1371/journal.pgen.1002964.

170. Tsofack SP, Meunier L, Sanchez L, Madore J, Provencher D, Mes-Masson A-M, Lebel M. Low expression of the X-linked ribosomal protein S4 in human serous epithelial ovarian cancer is associated with a poor prognosis. BMC Cancer. 2013;13(1):1–12.

171. Paquet ÉR, Hovington H, Brisson H, Lacombe C, Larue H, Têtu B, Lacombe L, Fradet Y, Lebel M. Low level of the X-linked ribosomal protein S4 in human urothelial carcinomas is associated with a poor prognosis. Biomark Med. 2015;9(3):187–97.

172. Zhou P, Xiang CX, Wei JF. The clinical significance of spondin 2 eccentric expression in peripheral blood mononuclear cells in bronchial asthma. J Clin Laby Anal. 2021;35(6):e23764.

173. Chang R, Chen L, Su G, Du L, Qin Y, Xu J, Tan H, Zhou C, Cao Q, Yuan G. Identification of ribosomal protein S4, Y-linked 1 as a cyclosporin a plus corticosteroid resistance gene. J Autoimmun. 2020;112:102465.

Publisher's Note

17

1.2 Publication Declaration - 'Sex-specific effects of *in utero* and adult tobacco smoke exposure'

Karosham D. Reddy & **Brian G.G. Oliver** (2020) *American Journal of Physiology – Lung Cellular and Molecular Physiology*.

Status: Published.

Author Contributions: Both authors conceived the ideas, read the manuscript and edited the final version. KDR drafted the manuscript. Both authors read and approved the final manuscript.

Signatures:

Name	Signature	Date
Karosham D. Reddy	Production Note: Signature removed prior to publication.	**08/12/2022**
Brian G.G. Oliver	Production Note: Signature removed prior to publication.	**08/12/2022**

MINI-REVIEW

Sex-specific effects of in utero and adult tobacco smoke exposure

Karosham D. Reddy[1,2] **and** Brian G. G. Oliver[1,2]

[1]*School of Life Sciences, University of Technology Sydney, Sydney, New South Wales, Australia; and* [2]*Respiratory Cellular and Molecular Biology, Woolcock Institute of Medical Research, University of Sydney, Sydney, New South Wales, Australia*

Abstract

Tobacco smoke has harmful effects on a multiorgan level. Exposure to smoke, whether in utero or environmental, significantly increases susceptibility. This susceptibility has been identified to be divergent between males and females. However, there remains a distinct lack of thorough research into the relationship between sex and exposure to tobacco. Females tend to generate a more significant response than males during adulthood exposure. The intrauterine environment is meticulously controlled, and exposure to tobacco presents a significant factor that contributes to poor health outcomes and susceptibility later in life. Analysis of these effects in relation to the sex of the offspring is yet to be holistically reviewed and summarized. In this review, we will delineate the time-dependent relationship between tobacco smoke exposure and sex-specific disease susceptibility. We further outline possible biological mechanisms that may contribute to the identified pattern.

disease susceptibility; in utero; mechanism; sexual dimorphism; tobacco

BACKGROUND

The intrauterine environment is highly regulated to enable healthy fetal development. Alterations to the molecular and chemical milieu predispose the fetus to chronic diseases later in life. The "developmental origins of disease" hypothesis describes that changes to prenatal conditions induce adaptions that alter anatomical structure and physiology to promote growth and success (26). These adaptations, although aiding intrauterine survival, may not offer protection perinatally (26). A recent review details the effects of maternal exposure to heavy metals, stress, tobacco smoke, and alcohol on the health outcomes of offspring (17), identifying that tobacco smoking negatively impacts multiple organ systems (17).

Tobacco smoking is a leading risk factor for noncommunicable diseases (NCDs). Although on the decline, smoking rates remain high worldwide. Smoke exposure remains the leading preventable cause of chronic diseases such as chronic obstructive pulmonary disease (COPD), cancer, and congenital heart disease (CHD) (3, 22, 36). Approximately 50% of women who smoke continue after discovering they are pregnant (45). As such, maternal tobacco smoke (MTS) is the leading preventable cause of abnormal pregnancy outcomes (45). MTS increases the likelihood of the newborn to suffer from various NCDs in childhood and adulthood (6, 42, 51, 60). The biological mechanism driving this is yet to be elucidated; however, hormones, epigenetics, oxidative stress, and DNA mutations have been postulated.

Active smoking has an overall adverse health impact, and sex-specific effects are recognized. Adult tobacco smoking has greater consequences in females, as they suffer higher morbidity and greater severity of tobacco-related disease compared with males (41). A meta-analysis found smoking intensity increased the magnitude and likelihood of stroke in females compared with males (50). Females had increased risk and severity of COPD and colorectal cancer than males with similar smoking levels (3, 22). Despite a clear pattern in adulthood, sex-specific prenatal effects of tobacco smoke exposure are rarely investigated appropriately or thoroughly.

This review aims to highlight the importance of fetal sex in response to tobacco smoke exposure in utero and compare these patterns to adult smoke exposure.

SEX DIFFERENCES IN DEVELOPMENT

Sexual dimorphism exists in utero, with the male and female fetus undergoing distinct physiological and developmental processes. This is apparent in lung development. Lung maturation is advanced in female fetuses, with mouth movement and surfactant production starting at 18 wk, whereas males reach this milestone after 20–23 wk (51, 54). As a result, the female fetus can generate a robust response to various exposures (54). The phenomenon of dysanapsis describes differences in fetal lung development between the sexes. Male lungs have larger but fewer airways, whereas females have more abundant yet smaller airways, enabling higher airflow rates (30). Increased flow rate enables greater adaptability in females to cope with adverse environments, sustaining stable growth. This phenomenon persists until 18 yr of age when male lung function overtakes females (30). The cause of delayed lung growth in males is unknown. It is a leading factor predisposing prepubertal males to asthma

19

and lower respiratory tract infections (LRTI) (66a), from which young females appear to be protected.

Sex differences in fetal brain, heart, and overall growth also exist. Genes located on the male-specific Y-chromosome drive sexual dimorphism in brain development (63). The Sex-determining Region Y (*SRY*) gene, primarily known for testicular maturation (63), controls a cascade of androgens regulating male brain development in utero (14). As evidence of dimorphism, boys are liable to the creation of white matter by androgen stimulation; on the other hand, estrogen in girls is inhibitory (14). Fetal heart rate patterns also differ, highlighting a divergence in male and female development. Males demonstrate increased heart rate variability (16), indicating a lack of stability in growth compared with females.

Here, we present inherent differences between the sexes in the absence of exogenous influences. However, the presence of this sexual dimorphism may leave either sex more susceptible to the effects of in utero exposure to external toxicants.

■ SEX-SPECIFIC SUSCEPTIBILITY DUE TO TOBACCO SMOKE

Modern medicine strives to design patient-specific "personalized medicine." The ideology is based on the heterogeneity of symptoms, pathology, and response to treatment. This concept explains why interventions are effective for some individuals, yet ineffective for others. As such, precise and distinct mechanisms regulate physiological processes in everyone. Although sex is a fundamental characteristic, it is widely overlooked or underinvestigated as a determining factor in disease and is considered a covariate or confounding factor. As such, limited thorough research exists investigating sex-specific effects of MTS and adult tobacco smoke exposure on disease susceptibility. Here, we will detail sex-specific pathological sensitivity due to both in utero and adult tobacco exposure.

Lung Function Decline

The process of dysanapsis is well described and closely linked to sex differences in respiratory outcomes (5, 74, 77). For example, males demonstrate decreased lung function compared with females who have a higher maximal expiratory flow at functional residual capacity (V'_{max} FRC) perinatally (74, 80). MTS compounds this sex difference in lung function, as shown in various cohort studies. A study of 6,740 Chinese children reported an independent negative correlation between MTS and postnatal smoke exposure on lung function, with a greater response in males overall. Male-adjusted odds ratio (OR) for decreased forced vital capacity (FVC) in response to MTS OR = 6.46 [95% confidence interval (CI) 2.58–16.17] compared with the females OR = 2.16 (95% CI 0.96–4.88) (33). Interestingly, the authors note an apparent protective effect of asthma diagnosis, with nonasthmatic offspring presenting a more pronounced decrease in lung function outcomes (33). A large North American cohort found that male MTS-exposed offspring consistently experience worse lung function outcomes [forced expiratory volume in one second (FEV_1), FEV_1/FVC, and reduced mid forced expiratory flow rates ($FEF_{25\%-75\%}$)] than female counterparts

when measured from 8 to 12 yr (15). Therefore, MTS has a lasting impact on lung development throughout childhood. The study is limited by its sample size of MTS-exposed children, most likely caused by underreporting of smoking by mothers, reducing the ability to differentiate MTS from postnatal smoke exposure. Nonetheless, a reduction in male offspring lung function from MTS is supported by multiple case-control cohort investigations (19, 29, 44, 49).

Chronic Obstructive Pulmonary Disease

COPD is primarily an adult-onset disease. It is characterized by accelerated lung function declines caused by prolonged exposure to noxious gases (68). As such, it is challenging to associate MTS to COPD incidence, as it is near impossible to differentiate prenatal influence from lifelong exposures. Generally, males report increased COPD incidence; however, evidence supports greater female susceptibility (25). A cross-sectional US cohort study concluded a significantly higher percentage (64%) of adults diagnosed with COPD were female from a cohort of 2,113 patients, possibly due to lifetime environmental tobacco exposure (22). The authors also report no significant relationship between MTS and COPD incidence ($P = 0.051$). Foreman et al. (2011) (25) find females predominate early-onset COPD due to adult exposure.

In contrast, a recent systematic review of 16 studies asserts that MTS may impair physiological development, reducing offspring lung function outcomes, and causing increased COPD risk in adulthood (65). Therefore, as there is a male-biased effect of MTS on lung function decline, men are predisposed toward COPD development later in life. However, the age window of susceptibility may vary between the sexes (69), requiring further investigation.

Asthma

The pattern of male sex being predominately affected by MTS disappears when investigating asthma. A sizable Australian cohort reported that when mothers smoke more than a pack a day, female offspring had an almost twofold increased risk of asthma; OR = 1.96 (95% CI 1.25–2.08) (1). Males showed no association between MTS and asthma at any age, with asthma risk reducing upon postnatal smoke exposure (1). A Finnish study investigated the effect of mothers smoking greater than and, less than 10 cigarettes during pregnancy, and identified no significant difference in asthma incidence between the sexes. A higher crude OR for female offspring of smoking mothers was reported, although males demonstrate the higher cumulative incidence of asthma irrespective of maternal smoking level (37). During prepubescence, males dominate asthma incidence. At puberty, this pattern reverses, and females dominate diagnoses with an increase in female cases rather than a decline in male patients, with disease incidence equalizing postmenopause (59). This natural pattern confounds many MTS asthma studies, which assess asthma diagnosis at age 14. As such, it is difficult to determine a definitive sex link between MTS and asthma. Thus, the methodology of future studies requires careful consideration to differentiate the temporal incidence of asthma from the effects of MTS. In general, women are more likely to use health care compared with men when matching for asthma status, indicating worse

disease prognosis (41). Asthmatic females have greater bronchial hyperresponsiveness to cigarette smoke, highlighting another sexual dimorphism (41). In support, a large Swedish cohort found smoking increases the female relative risk of adult-onset asthma with an increased rate ratio of 1.3 (95% CI 1.0–1.6) compared with males (72). Therefore, the respiratory effects of adult smoke exposure are more pronounced in females.

Multiorgan Effects

Sexual dimorphism in disease susceptibility due to tobacco smoke exposure extends beyond the respiratory system, having multiorgan disease influences. Childhood cancers such as leukemia, childhood brain tumors (CBT), and lymphomas report higher incidence in boys than girls (6). Adult female smokers have a greater rate of cancer, despite lower smoking intensity than males (3). Anderson et al. (2010) (3) found a statistically significant increased risk of advanced colorectal neoplasia for mild smoking (10–30 pack-years) females OR = 4.11 (95% CI 1.88–9.01). Males only show increased disease development risk with heavy smoking (≥30 pack-years) OR = 3.10 (95% CI 1.71–5.65) that is similar to females at this smoking intensity. This contrasts with prenatal exposure that is linked to increased risk of disease in males. Tettamanti et al. (2016) (71) identified in a large Swedish cohort that smoking during pregnancy indicated a trend toward increased childhood brain tumor incidence in male children aged between 5 and 9 yr old, with nonsignificant OR = 1.23 (95% CI 0.91–1.67). A recent review supports this trend asserting that nicotine affects male brain development more than in females (14). However, in contrast, large cohort studies find either a slight increase or no link between MTS and the development of CBT (27, 76).

Intellectual and social disability is linked to the action of nicotine on fetal brain development (14). Meta-analyses and cohort studies find increased male susceptibility to intellectual disability (34, 35). MTS is asserted to primarily affect brain microstructure in males, causing less-coherent fibers and myelination (13). This male-dominated risk is attributed to genetic differences defined by the Y-chromosome (not present in females), resulting in differential expression of genes in specific brain regions (63).

Smoke exposure increases congenital heart disease (CHD) risk with teratogenic effects during early pregnancy fetal heart development. The incidence of CHD is higher in children of smokers (60.9%) compared with nonsmokers (35.8%) (39). A study of 365 neonates found an increased OR = 2.75 (95% CI 1.66–4.58) of CHD due to periconceptual smoking, increasing with smoking intensity (39). Thus, the effects of tobacco smoke are highly persistent, possibly affecting gametes. A Brazilian longitudinal cohort found female offspring exposed to MTS presented decreased HDL levels (31). Although an indirect indicator, the authors assert females may be more sensitive to cardiovascular disease due to MTS because of reduced circulating high-density lipoprotein (HDL). Another study found that the female offspring of smokers had lower HDL levels than males in a large British cohort (60). These findings do not account for the pediatric effects of MTS, limiting the ability to make conclusions regarding the lifelong effects of smoke exposure. Adult

female smokers have an overall increased risk of cardiovascular disease, demonstrating a 50% increased risk of myocardial infarction due to smoking (36, 61). Male mice offspring exposed to prenatal nicotine demonstrated a higher risk of hypertension and increased arterial thickness, with no significant effect observed for females (79). Alexander et al. (2011) (2) assert that precise timing of in utero exposures significantly impacts the subsequent cardiovascular outcome, accounting for significant variability in the findings of the literature relating to cardiovascular disease.

SEX-SPECIFIC SUSCEPTIBILITY DUE TO E-CIGARETTE EXPOSURE

The recent popularity of e-cigarettes or vapes as an alternative to traditional tobacco smoking is vital to acknowledge. Due to their perception of being "safer," e-cigarette usage in pregnancy equals tobacco cigarette consumption (78). Limited studies of the long-term physiological effects of e-cigarette products exist. McGrath-Morrow et al. (2015) (46) identify a modest impairment of lung growth caused by neonatal e-cigarette exposure but do not investigate any effect of sex. An epidemiological study reported males comprise ~69% of patients suffering from lung injury associated with vaping, indicating a greater pathogenic effect in males (57). However, this was not normalized for vaping frequency in men and women. As there are striking adverse effects of e-cigarettes responses in vitro (8), and in animal models (53), further research into the long-term and sex-specific effects of e-cigarettes is needed.

Across all NCDs discussed, the effect of MTS was weakened with age (11), indicating a more substantial impact on pediatric outcomes and less of a lifelong impact. We have identified that prenatal smoke exposure predominately affects males, with a higher incidence and susceptibility of disease states across multiple organ systems. Some studies report MTS-related effects in females; however, this is confounded by unaccounted postnatal environmental and biological factors. The male fetus appears to be more susceptible to the adverse effects of tobacco smoke, whereas female fetuses are protected. This protection only lasts for the childhood years, with adult females consistently experiencing more extreme forms of chronic disease from tobacco smoke exposure (3, 12, 74). Our findings are summarized in Table 1.

MECHANISMS OF SUSCEPTIBILITY

Cleary, sexual dimorphism exists in disease susceptibility due to tobacco smoke exposure. Importantly, the affected sex differs depending on in utero versus adulthood exposure. Here, we outline possible mechanisms that may contribute to sex differences.

Sex Steroid Hormones

Hormones are a critical biological factor responsible for various phenotypic differences between males and females. The primary male hormones are androgens (testosterone). At the same time, in females, estrogens (estradiol) predominate. However, both function in either sex. Estrogens increase the production of proinflammatory cytokines such as TNF-α,

21

Table 1. *Summary of sexual dimorphism in disease susceptibility due to tobacco smoke exposure prenatally and during adulthood*

Condition	Evidence	Tobacco Smoke Exposure Time	
		Prenatal	Adult
Cancer	Tettamanti et al. (2016) (71)	Male	Male
	Botsivali and Kyrtopoulos (2019) (6)	Male	
	Anderson et al. (2011) (3)		Female
	Drake et al. (2015) (21)	Male	
Respiratory	Jaakola and Gissler (2007) (37)	Female	
	Dotterud et al. (2013) (20)	Male	
	Hu et al. (2017) (33)	Male	
	Cunningham, Dockery, and Spelzer (1994) (15)	Male	
	Hayatbakhsh et al. (2009) (29)	Male	
	Torén and Hermansson (1999) (72)		Female
	Kynyk, Mastronarde, and McCallister (2011) (41)		Female
	Miller et al. (2014) (48)	Female	
	Moshammer et al. (2006) (49)	Male	
	Eisner et al. (2005) (22)		Female
	Foreman et al. (2011) (25)		Female
Neurological conditions	Hutchinson et al. (2009) (35)	Male	
	Cross, Linker, and Leslie (2017) (14)	Male	
	Appelman et al. (2015) (4)		Female
	Peters, Huxley, and Woodward (2013) (58)		Female
	Chang et al. (2016) (13)	Male	
Cardiovascular diseases	Power, Atherton, and Thomas (2010) (60)	Female	
	Horta et al. (2011) (31)	Female	
	Prescott et al. (1998) (61)		Female
	Huxley and Woodward (2011) (36)		Female
	Xiao et al. (2008) (79)	Male	

IFN-γ, and IL-12 while downregulating the anti-inflammatory cytokine IL-10 (66). Androgens have the opposite effect, increasing IL-10 levels and suppressing the production of IFN-γ (66). Increased estrogen levels correlate with greater production of cytochrome P450 enzymes that are central to the metabolism of tobacco into toxic intermediaries (70). As a result, there is accelerated production of oxidative smoke metabolites that contribute to lung injury and inflammation. The prevalence and strength of the estrogen's effects change over time. Mature young females generate a more robust immune response, compared with males and older women, which is attributed to higher blood estrogen levels (56). The immune response in male mice is increased with exogenous estrogen, indicating a similar function of the hormone in both sexes (47, 66).

Conversely, androgens act to suppress the innate immune response upregulating IL-10 production. Maternal testosterone levels are 11% higher in smokers, whereas no effect is seen for estrogen (73). Generally, postmenopausal women tend to respond similarly to external stimuli compared with males; this is thought to be related to estrogen levels declining with age. However, it remains an understudied phenomenon. Brand et al. (2011) (9) found smoking increases circulating levels of both male- and female-associated hormones in postmenopausal women, but these were not comparable to perimenopause levels. Therefore, a close relationship exists between tobacco smoke exposure and sex hormone levels in circulation.

Epigenetics

Males and females have similar genomes, except for the male-specific Y-chromosome; therefore, sex-biased gene regulation is likely occurring. Our group has previously shown that sexual dimorphism is maintained ex vivo (62), supporting the notion that an internal mechanism other than hormones

drives observed sex differences. It is well established that MTS alters the epigenome causing adverse effects in the lung (81). Epigenetics is a mechanism influencing transcription, showing clear sexual-dimorphism in utero (43, 51). Ladd-Acosta et al. (2016) (43) describe MTS-induced sex-specific methylation patterns, where MTS-exposed female preschoolers had increased methylation, with decreased methylation in males near the *HLA-DPB2* gene locus. Other studies identify that genes involved in the immune response and defense against infection are altered in female infants. Genes that regulate cancer, developmental, and neurological processes are affected in male infants (40). This partially explains why males incur more developmental problems, particularly in childhood, while women experienced more smoking-induced chronic inflammatory diseases in adulthood. Many smoking-related genes are differentially methylated in utero, for example, the *AHRR* and *CYP1A1* genes are central in tobacco smoke metabolism, and *GFI1*, which is essential in developmental processes and histone modification (38). Significantly, differential methylation caused by smoking shows no differences between males and females in whole blood (82). Upon more in-depth investigation, the authors identify that males have an increased tendency for altered DNA methylation, with 42 differentially methylated sites present in males compared with 10 locations in females. Females demonstrate alteration to immune-related genes due to adult smoke exposure, such as *CCL13*, *CCL2*, and *FCN1* (18). Breitling et al. (2011) (10) assert that sexual dimorphism of DNA methylation does not occur due to adulthood tobacco smoke exposure. Figure 1 highlights that this dynamic relationship between MTS and gene expression in utero presents as highly valuable future research pathways to generate a holistic picture between sex, the epigenome, and tobacco smoke.

Female Adult Smoke Exposure

- Limited sex difference in methylation patterns
- ♀ Altered immune-related genes
  - *CD4*
  - *CCL13*
  - *CCL2*
  - *FCN1*

Epigenetic Changes

Me

Me

Male Adult Smoke Exposure

- Limited sex difference in methylation patterns
- ♂ Slight increased susceptibility to alterations in DNA methylation by smoking

Maternal Tobacco Smoke Exposure
- ♀ Increased gene methylation in females
- ♂ Decreased gene methylation in males
- ♂ DNA methylation aberrations more induced in males
- Various sex differences in methylation patterns at birth

Figure 1. Sexual dimorphism in epigenetic changes between males and females depending on the time of tobacco smoke exposure. ♂ = male-specific effect, ♀ = female-specific effect.

Telomere Length

Telomeres are complicated nucleotide sequences located at the end of chromosomes and functioning in cell division, chromosome stabilization, and apoptosis (64). Telomere length (TL) contributes to a variety of diseases. Shorter TL is linked to atherosclerosis and cardio-vascular diseases, whereas long TL is associated with cancers as cellular longevity is increased (24). An association between the sex-specific relationship between telomere length and smoking has been identified and is summarized in Fig. 2. MTS is linked to shorter TL in males (5a); however, studies also find females have longer TL than males across all ages, confounding these findings (24). An

Adult Females

- ♀ Longer telomere length
- ⬆ Increased resilience to smoke exposure
- ⬆ Increased risk of cancer

Telomere Length

Adult Males

- ♂ Shorter telomere length
- ⬇ Reduced resilience to tobacco exposure
- ⬆ Increased cardiovascular disease risk

Maternal Tobacco Smoke Exposure
- ♂ Shortened telomere length in males
- ⬇ Shortened life expectancy
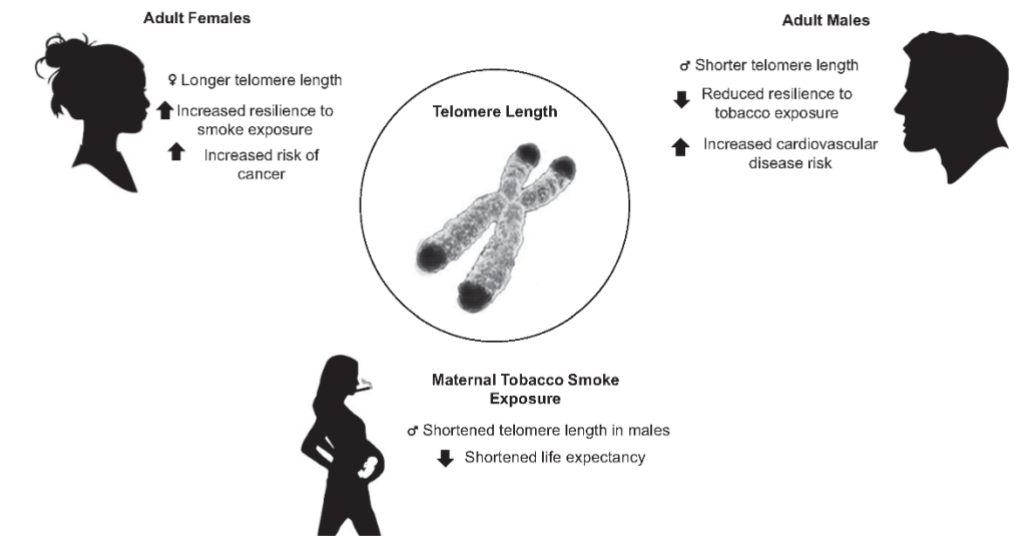
Figure 2. Sexual dimorphism in the effect of tobacco smoke exposure on telomere length depending on the time of exposure between males and females. ♂ = male-specific effect, ♀ = female-specific effect.

23

Adult Females

♀ Increased female mitochondrial activity

▲ Greater resilience to environmental stressors

▼ Decreased risk of metabolic disorders

Mitochondrial Response

Adult Males

♂ Reduced male mitochondrial activity

▼ Decreased antioxidant capacity

▲ Increased production of ROS

▲ Increased risk of DNA damage/mutation

Maternal Tobacco Smoke Exposure

▲ Mitochondrial DNA number

♂ Breakdown of mitochondrial function in males

▲ Drp-1

▲ Opa-1

♀ No effect in females

Figure 3. Sexual dimorphism in mitochondrial responses between adult males and females, and sex-specific effects of prenatal tobacco smoke exposure on fetal mitochondria. ♂ = male-specific effect, ♀ = female-specific effect.

in vitro investigation of telomeres in cord blood found only male telomere length was negatively affected by smoke exposure in utero (5a). The generally longer TL in females may enable more stability and resilience to exposures, subsequently offering protection from short TL-associated diseases postnatally (28). Therefore, it is pertinent to investigate how smoking directly influences TL and its link to sex-specific disease susceptibility.

Telomere Length

♂ Shortened telomere length in males

Mitochondrial Response

♂ Breakdown of mitochondrial function in males

▲ Drp-1

▲ Opa-1

♀ No effect seen in females

Epigenetics

♀ Increased methylation of immuno-regulatory genes

♂ Decreased methylation of developmental and neurological genes

Maternal Tobacco Smoke Exposure Mechanisms

Testosterone

▲ Higher levels in smoking mothers

▼ Reduced immune response in offspring

Estrogen

- No change with maternal smoking

Figure 4. Schematic summarizing molecular mechanisms contributing to sex differences in disease susceptibility of offspring in response to maternal tobacco smoke exposure. ♂ = male-specific effect, ♀ = female-specific effect.

24

Mitochondria

Mitochondria produce the energy currency of the cell, namely, adenosine-tri-phosphate (ATP). But mitochondrial functions extend beyond ATP production from hormone synthesis to ionic regulation and apoptosis (75). The breakdown of mitochondria bioenergetics in utero affects various biological processes, proving detrimental to fetal health. Tobacco smoke increases mitochondrial DNA copy number to compensate for a decline in cellular respiratory function (43a). MTS induces a breakdown in enzymatic activity of placental mitochondria (7), negatively impacting the ability to sustain optimum intrauterine conditions. Toxic products in tobacco smoke accumulate in the placenta (7), with tobacco smoke metabolized at a higher rate in the fetus. Therefore, the interaction between smoking by-products such as CO, nicotine, and thiocyanate contributes to a breakdown in mitochondrial function (7). Our group recently found that male mice offspring exposed to smoke in utero showed an increase in mitophagy markers [Drp-1 and Opa-1, while no effect was seen in female mice (55)]. These proteins are markers of mitochondrial fission and fusion, indicating that mitochondrial renewal is occurring. Dysregulation of these processes results in dysfunctional mitochondria that correlate with disease severity in asthma and COPD. The optimized function of mitochondria in females was recognized in peripheral blood mononuclear cells, where significantly greater mitochondrial activity was reported in women (67). It is hypothesized that as mitochondria are maternally inherited, spending more time under selection in females, they are optimized for function in females rather than males. Due to greater functional ability, female mitochondria may have more resilience and improved ability to respond to stressors such as environmental smoke, thus protecting females from reactive oxygen species (ROS) and oxidative stress. Ventura-Clapier et al. (2017) (75) thoroughly detail the tissue and sex specificity of mitochondria. Figure 3 illustrates temporal sex differences of mitochondrial response, highlighting that the mitochondria may significantly contribute to the disparity in disease susceptibility between the sexes and warrants more in-depth investigation.

CONCLUSIONS

The intrauterine environment is a highly regulated and dynamic milieu of factors. Hence, exposure of the fetus to tobacco smoke threatens health outcomes for the offspring. We have reviewed sexual dimorphism in disease susceptibility depending on the time line of exposure to tobacco smoke. We found compounding evidence that male offspring exposed to cigarette smoke in utero have increased sensitivity to disease development in multiple organ systems. This trend is reversed when exposure takes place in adulthood, where females show greater susceptibility (4). Therefore, a distinct sexual dimorphism exists regarding the effects of tobacco smoke exposure, with the time line of exposure having a defining impact, as illustrated in Figs. 4 5.

We have also discussed possible biological mechanisms driving sexual dimorphism. Sex hormones, epigenetics, telomere length, and the mitochondrial response may contribute either independently or concurrently to the sex disparity. Other possible mechanisms contributing to this
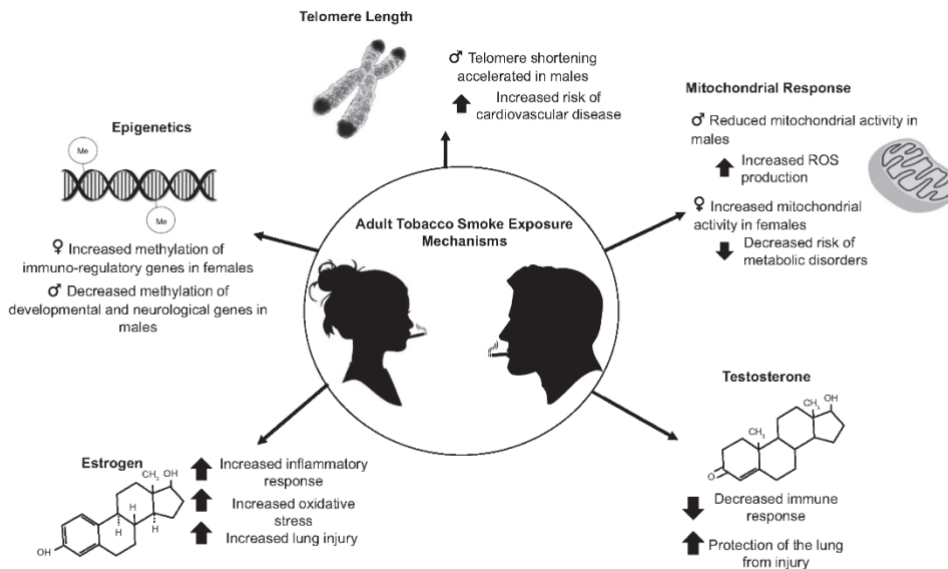
Figure 5. Schematic summarizing molecular mechanisms contributing to sex differences in disease susceptibility caused by adulthood tobacco smoke exposure. ♂ = male-specific effect, ♀ = female-specific effect.

pattern that are not discussed in this review include genetic polymorphisms, chromosomal aberrations, and the sex chromosomes. As a distinct temporal switch exists in sex susceptibility, it is possible that different mechanisms function at different life stages. Future research must focus on the timing and intensity of exposures to develop a holistic understanding. The outcomes of subsequent mechanistic studies can inform clinical responses, enabling patient-specific interventions, and more effective treatments.

GRANTS

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

K.D.R. and B.G.G.O. conceived and designed review; drafted manuscript; edited and revised manuscript; and approved final version of manuscript.

REFERENCES

1. **Alati R, Al Mamun A, O'Callaghan M, Najman JM, Williams GM.** In utero and postnatal maternal smoking and asthma in adolescence. *Epidemiology* 17: 138–144, 2006. doi:10.1097/01.ede.0000198148.02347.33.

2. **Alexander BT, Dasinger JH, Intapad S.** Fetal programming and cardiovascular pathology. *Compr Physiol* 5: 997–1025, 2015. doi:10.1002/cphy.c140036.

3. **Anderson JC, Moezardalan K, Messina CR, Latreille M, Shaw RD.** Smoking and the association of advanced colorectal neoplasia in an asymptomatic average risk population: analysis of exposure and anatomical location in men and women. *Dig Dis Sci* 56: 3616–3623, 2011. doi:10.1007/s10620-011-1814-8.

4. **Appelman Y, van Rijn BB, Ten Haaf ME, Boersma E, Peters SA.** Sex differences in cardiovascular risk factors and disease prevention. *Atherosclerosis* 241: 211–218, 2015. doi:10.1016/j.atherosclerosis.2015.01.027.

5. **Becklake MR, Kauffmann F.** Gender differences in airway behaviour over the human life span. *Thorax* 54: 1119–1138, 1999. doi:10.1136/thx.54.12.1119.

5a. **Bosquet Enlow M, Bollati V, Sideridis G, Flom JD, Hoxha M, Hacker MR, Wright RJ.** Sex differences in effects of maternal risk and protective factors in childhood and pregnancy on newborn telomere length. *Psychoneuroendocrinology* 95: 74–85, 2018. doi:10.1016/j.psyneuen.2018.05.025.

6. **Botsivali M, Kyrtopoulos SA.** Transplacental exposure to carcinogens and risks to children: evidence from biomarker studies and the utility of omic profiling. *Arch Toxicol* 93: 833–857, 2019. doi:10.1007/s00204-019-02428-3.

7. **Bouhours-Nouet N, May-Panloup P, Coutant R, de Casson FB, Descamps P, Douay O, Reynier P, Ritz P, Malthièry Y, Simard G.** Maternal smoking is associated with mitochondrial DNA depletion and respiratory chain complex III deficiency in placenta. *Am J Physiol Endocrinol Metab* 288: E171–E177, 2005. doi:10.1152/ajpendo.00260.2003.

8. **Bozier J, Rutting S, Xenaki D, Peters M, Adcock I, Oliver BG.** Heightened response to e-cigarettes in COPD. *ERJ Open Res* 5: 00192–02018, 2019. doi:10.1183/23120541.00192-2018.

9. **Brand JS, Chan M-F, Dowsett M, Folkerd E, Wareham NJ, Luben RN, van der Schouw YT, Khaw K-T.** Cigarette smoking and endogenous sex hormones in postmenopausal women. *J Clin Endocrinol Metab* 96: 3184–3192, 2011. doi:10.1210/jc.2011-1165.

10. **Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H.** Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. *Am J Hum Genet* 88: 450–457, 2011. doi:10.1016/j.ajhg.2011.03.003.

11. **Burke H, Leonardi-Bee J, Hashim A, Pine-Abata H, Chen Y, Cook DG, Britton JR, McKeever TM.** Prenatal and passive smoke exposure and incidence of asthma and wheeze: systematic review and meta-analysis. *Pediatrics* 129: 735–744, 2012. doi:10.1542/peds.2011-2196.

12. **Casimir GJ, Lefèvre N, Corazza F, Duchateau J.** Sex and inflammation in respiratory diseases: a clinical viewpoint. *Biol Sex Differ* 4: 16, 2013. doi:10.1186/2042-6410-4-16.

13. **Chang L, Oishi K, Skranes J, Buchthal S, Cunningham E, Yamakawa R, Hayama S, Jiang CS, Alicata D, Hernandez A, Cloak C, Wright T, Ernst T.** Sex-specific alterations of white matter developmental trajectories in infants with prenatal exposure to methamphetamine and tobacco. *JAMA Psychiatry* 73: 1217–1227, 2016. doi:10.1001/jamapsychiatry.2016.2794.

14. **Cross SJ, Linker KE, Leslie FM.** Sex-dependent effects of nicotine on the developing brain. *J Neurosci Res* 95: 422–436, 2017. doi:10.1002/jnr.23878.

15. **Cunningham J, Dockery DW, Speizer FE.** Maternal smoking during pregnancy as a predictor of lung function in children. *Am J Epidemiol* 139: 1139–1152, 1994. doi:10.1093/oxfordjournals.aje.a116961.

16. **DiPietro JA, Costigan KA, Voegtline KM.** Studies in fetal behavior: revisited, renewed, and reimagined. *Monogr Soc Res Child Dev* 80: vii, 2015. doi:10.1111/mono.12170.

17. **DiPietro JA, Voegtline KM.** The gestational foundation of sex differences in development and vulnerability. *Neuroscience* 342: 4–20, 2017. doi:10.1016/j.neuroscience.2015.07.068.

18. **Dogan MV, Shields B, Cutrona C, Gao L, Gibbons FX, Simons R, Monick M, Brody GH, Tan K, Beach SR, Philibert RA.** The effect of smoking on DNA methylation of peripheral blood mononuclear cells from African American women. *BMC Genomics* 15: 151, 2014. doi:10.1186/1471-2164-15-151.

19. **Dong G-H, Wang D, Yang Z-H, Zhang P-F, Ren W-H, Zhao Y-D, He Q-C.** Gender-specific differences in effects of prenatal and postnatal environmental tobacco smoke exposure on respiratory symptoms in 23,474 children with and without allergic predisposition: results from 25 districts of northeast China. *Int J Environ Health Res* 21: 173–188, 2011. doi:10.1080/09603123.2010.515673.

20. **Dotterud CK, Storrø O, Simpson MR, Johnsen R, Øien T.** The impact of pre- and postnatal exposures on allergy related diseases in childhood: a controlled multicentre intervention study in primary health care. *BMC Public Health* 13: 123, 2013. doi:10.1186/1471-2458-13-123.

21. **Drake AJ, O'Shaughnessy PJ, Bhattacharya S, Monteiro A, Kerrigan D, Goetz S, Raab A, Rhind SM, Sinclair KD, Meharg AA, Feldmann J, Fowler PA.** In utero exposure to cigarette chemicals induces sex-specific disruption of one-carbon metabolism and DNA methylation in the human fetal liver. *BMC Med* 13: 18, 2015. doi:10.1186/s12916-014-0251-x.

22. **Eisner MD, Balmes J, Katz PP, Trupin L, Yelin EH, Blanc PD.** Lifetime environmental tobacco smoke exposure and the risk of chronic obstructive pulmonary disease. *Environ Health* 4: 7, 2005. doi:10.1186/1476-069X-4-7.

24. **Factor-Litvak P, Susser E, Kezios K, McKeague I, Kark JD, Hoffman M, Kimura M, Wapner R, Aviv A.** Leukocyte telomere length in newborns: implications for the role of telomeres in human disease. *Pediatrics* 137: e20153927, 2016. doi:10.1542/peds.2015-3927.

25. **Foreman MG, Zhang L, Murphy J, Hansel NN, Make B, Hokanson JE, Washko G, Regan EA, Crapo JD, Silverman EK, DeMeo DL; COPDGene Investigators.** Early-onset chronic obstructive pulmonary disease is associated with female sex, maternal factors, and African American race in the COPDGene Study. *Am J Respir Crit Care Med* 184: 414–420, 2011. doi:10.1164/rccm.201011-1928OC.

26. **Godfrey KM, Barker DJ.** Fetal programming and adult health. *Public Health Nutr* 4, 2B: 611–624, 2001. doi:10.1079/PHN2001145.

27. **Greenop KR, Blair EM, Bower C, Armstrong BK, Milne E.** Factors relating to pregnancy and birth and the risk of childhood brain tumors: results from an Australian case-control study. *Pediatr Blood Cancer* 61: 493–498, 2014. doi:10.1002/pbc.24751.

26

28. **Hallows SE, Regnault TR, Betts DH.** The long and short of it: the role of telomeres in fetal origins of adult disease. *J Pregnancy* 2012: 638476, 2012. doi:10.1155/2012/638476.

29. **Hayatbakhsh MR, Sadasivam S, Mamun AA, Najman JM, Williams GM, O'Callaghan MJ.** Maternal smoking during and after pregnancy and lung function in early adulthood: a prospective study. *Thorax* 64: 810–814, 2009. doi:10.1136/thx.2009.116301.

30. **Hibbert M, Lannigan A, Raven J, Landau L, Phelan P.** Gender differences in lung growth. *Pediatr Pulmonol* 19: 129–134, 1995. doi:10.1002/ppul.1950190208.

31. **Horta BL, Gigante DP, Nazmi A, Silveira VMF, Oliveira I, Victora CG.** Maternal smoking during pregnancy and risk factors for cardiovascular disease in adulthood. *Atherosclerosis* 219: 815–820, 2011. doi:10.1016/j.atherosclerosis.2011.08.018.

33. **Hu L-W, Yang M, Chen S, Shah K, Hailegiorgis Y, Burgens R, Vaughn M, Huang J, Xaverius P, Paul G, Morawska L, Lu T, Lin S, Zhong SQ, Kong ML, Xie YQ, Hao YT, Zeng XW, Qian Z, Dong GH.** Effects of in utero and post-natal exposure to secondhand smoke on lung function by gender and asthma status: the Seven Northeastern Cities (SNEC) Study. *Respiration* 93: 189–197, 2017. doi:10.1159/000455140.

34. **Huang J, Zhu T, Qu Y, Mu D.** Prenatal, perinatal and neonatal risk factors for intellectual disability: a systemic review and meta-analysis. *PLoS One* 11: e0153655, 2016. doi:10.1371/journal.pone.0153655.

35. **Hutchinson J, Pickett KE, Green J, Wakschlag LS.** Smoking in pregnancy and disruptive behaviour in 3-year-old boys and girls: an analysis of the UK Millennium Cohort Study. *J Epidemiol Community Health* 64: 82–88, 2010. doi:10.1136/jech.2009.089334.

36. **Huxley RR, Woodward M.** Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies. *Lancet* 378: 1297–1305, 2011. doi:10.1016/S0140-6736(11)60781-2.

37. **Jaakkola JJ, Gissler M.** Are girls more susceptible to the effects of prenatal exposure to tobacco smoke on asthma? *Epidemiology* 18: 573–576, 2007. doi:10.1097/EDE.0b013e31812001d2.

38. **Joubert BR, Håberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, Huang Z, Hoyo C, Midttun Ø, Cupul-Uicab LA, Ueland PM, Wu MC, Nystad W, Bell DA, Peddada SD, London SJ.** 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect* 120: 1425–1431, 2012. doi:10.1289/ehp.1205412.

39. **Karatza AA, Giannakopoulos I, Dassios TG, Belavgenis G, Mantagos SP, Varvarigou AA.** Periconceptional tobacco smoking and isolated congenital heart defects in the neonatal period. *Int J Cardiol* 148: 295–299, 2011. doi:10.1016/j.ijcard.2009.11.008.

40. **Khulan B, Cooper WN, Skinner BM, Bauer J, Owens S, Prentice AM, Belteki G, Constancia M, Dunger D, Affara NA.** Periconceptional maternal micronutrient supplementation is associated with widespread gender related changes in the epigenome: a study of a unique resource in the Gambia. *Hum Mol Genet* 21: 2086–2101, 2012. doi:10.1093/hmg/dds026.

41. **Kynyk JA, Mastronarde JG, McCallister JW.** Asthma, the sex difference. *Curr Opin Pulm Med* 17: 6–11, 2011. doi:10.1097/MCP.0b013e3283410038.

42. **La Merrill MA, Cirillo PM, Krigbaum NY, Cohn BA.** The impact of prenatal parental tobacco smoking on risk of diabetes mellitus in middle-aged women. *J Dev Orig Health Dis* 6: 242–249, 2015. doi:10.1017/S2040174415000045.

43. **Ladd-Acosta C, Shu C, Lee BK, Gidaya N, Singer A, Schieve LA, Schendel DE, Jones N, Daniels JL, Windham GC, Newschaffer CJ, Croen LA, Feinberg AP, Daniele Fallin M.** Presence of an epigenetic signature of prenatal cigarette smoke exposure in childhood. *Environ Res* 144, Pt A: 139–148, 2016. doi:10.1016/j.envres.2015.11.014.

43a. **Lee HC, Yin PH, Lu CY, Chi CW, Wei YH.** Increase of mitochondria and mitochondrial DNA in response to oxidative stress in human cells. *Biochem J* 348: 425–432, 2000. doi:10.1042/bj3480425.

44. **Li Y-F, Gilliland FD, Berhane K, McConnell R, Gauderman WJ, Rappaport EB, Peters JM.** Effects of in utero and environmental tobacco smoke exposure on lung function in boys and girls with and without asthma. *Am J Respir Crit Care Med* 162: 2097–2104, 2000. doi:10.1164/ajrccm.162.6.2004178.

45. **McEvoy CT, Spindel ER.** Pulmonary effects of maternal smoking on the fetus and child: effects on lung development, respiratory morbidities, and life long lung health. *Paediatr Respir Rev* 21: 27–33, 2017. doi:10.1016/j.prrv.2016.08.005.

46. **McGrath-Morrow SA, Hayashi M, Aherrera A, Lopez A, Malinina A, Collaco JM, Neptune E, Klein JD, Winickoff JP, Breysse P, Lazarus P, Chen G.** The effects of electronic cigarette emissions on systemic cotinine levels, weight and postnatal lung growth in neonatal mice. *PLoS One* 10: e0118344, 2015. doi:10.1371/journal.pone.0118344.

47. **Milette S, Hashimoto M, Perrino S, Qi S, Chen M, Ham B, Wang N, Istomine R, Lowy AM, Piccirillo CA, Brodt P.** Sexual dimorphism and the role of estrogen in the immune microenvironment of liver metastases. *Nat Commun* 10: 5745, 2019. doi:10.1038/s41467-019-13571-x.

48. **Miller LL, Henderson J, Northstone K, Pembrey M, Golding J.** Do grandmaternal smoking patterns influence the etiology of childhood asthma? *Chest* 145: 1213–1218, 2014. doi:10.1378/chest.13-1371.

49. **Moshammer H, Hoek G, Luttmann-Gibson H, Neuberger MA, Antova T, Gehring U, Hruba F, Pattenden S, Rudnai P, Slachtova H, Zlotkowska R, Fletcher T.** Parental smoking and lung function in children: an international study. *Am J Respir Crit Care Med* 173: 1255–1263, 2006. doi:10.1164/rccm.200510-1552OC.

50. **Mucha L, Stephenson J, Morandi N, Dirani R.** Meta-analysis of disease risk associated with smoking, by gender and intensity of smoking. *Gend Med* 3: 279–291, 2006. doi:10.1016/S1550-8579(06)80216-0.

51. **Murphy SK, Adigun A, Huang Z, Overcash F, Wang F, Jirtle RL, Schildkraut JM, Murtha AP, Iversen ES, Hoyo C.** Gender-specific methylation differences in relation to prenatal exposure to cigarette smoke. *Gene* 494: 36–43, 2012. doi:10.1016/j.gene.2011.11.062.

53. **Nguyen T, Li GE, Chen H, Cranfield CG, McGrath KC, Gorrie CA.** Maternal E-cigarette exposure results in cognitive and epigenetic alterations in offspring in a mouse model. *Chem Res Toxicol* 31: 601–611, 2018. doi:10.1021/acs.chemrestox.8b00084.

54. **Noël A, Xiao R, Perveen Z, Zaman H, Le Donne V, Penn A.** Sex-specific lung functional changes in adult mice exposed only to secondhand smoke in utero. *Respir Res* 18: 104, 2017. doi:10.1186/s12931-017-0591-0.

55. **Oliver B, Wang B, Chan Y, Zhou S, Saad S, Chen H.** Offspring sex affects the susceptibility to maternal smoking-induced lung inflammation and the effect of maternal antioxidant supplementation in mice [Preprint]. *Research Square*, 2020. doi:10.21203/rs.2.19269/v2+.

56. **Pardue M-L, Wizemann TM.** *Exploring the Biological Contributions to Human Health: Does Sex Matter?* Washington, DC: National Academies Press, 2001.

57. **Perrine CG, Pickens CM, Boehmer TK, King BA, Jones CM, DeSisto CL,** et al. Characteristics of a multistate outbreak of lung injury associated with e-cigarette use, or vaping—United States, 2019. *MMWR Morb Mortal Wkly Rep* 68: 860–864, 2019. doi:10.15585/mmwr.mm6839e1.

58. **Peters SA, Huxley RR, Woodward M.** Smoking as a risk factor for stroke in women compared with men: A systematic review and meta-analysis of 81 cohorts, including 3 980 359 individuals and 42 401 strokes. *Stroke* 44: 2821–2828, 2013. doi:10.1161/STROKEAHA.113.002342.

59. **Postma DS.** Gender differences in asthma development and progression. *Gend Med* 4, Suppl B: S133–S146, 2007. doi:10.1016/S1550-8579(07)80054-4.

60. **Power C, Atherton K, Thomas C.** Maternal smoking in pregnancy, adult adiposity and other risk factors for cardiovascular disease. *Atherosclerosis* 211: 643–648, 2010. doi:10.1016/j.atherosclerosis.2010.03.015.

61. **Prescott E, Hippe M, Schnohr P, Hein HO, Vestbo J.** Smoking and risk of myocardial infarction in women and men: longitudinal population study. *BMJ* 316: 1043–1047, 1998. doi:10.1136/bmj.316.7137.1043.

62. **Reddy KD, Rutting S, Tonga K, Xenaki D, Simpson JL, McDonald VM, Plit M, Malouf M, Zakarya R, Oliver BG.** Sexually dimorphic production of interleukin-6 in respiratory disease. *Physiol Rep* 8: e14459, 2020. doi:10.14814/phy2.14459.

63. **Reinius B, Jazin E.** Prenatal sex differences in the human brain. *Mol Psychiatry* 14: 987, 2009. doi:10.1038/mp.2009.79.

64. **Salihu HM, Pradhan A, King L, Paothong A, Nwoga C, Marty PJ, Whiteman V.** Impact of intrauterine tobacco exposure on fetal telomere length. *Am J Obstet Gynecol* 212: 205.e1–205.e8, 2015. doi:10.1016/j.ajog.2014.08.026.

27

65. **Savran O, Ulrik CS.** Early life insults as determinants of chronic obstructive pulmonary disease in adult life. *Int J Chron Obstruct Pulmon Dis* 13: 683–693, 2018. doi:10.2147/COPD.S153555.

66. **Schurz H, Salie M, Tromp G, Hoal EG, Kinnear CJ, Möller M.** The X chromosome and sex-specific effects in infectious disease susceptibility. *Hum Genomics* 13: 2, 2019. doi:10.1186/s40246-018-0185-z.

66a. **Shah R, Newcomb DC.** Sex bias in asthma prevalence and pathogenesis. *Front Immunol* 9: 2997, 2018. doi:10.3389/fimmu.2018.02997.

67. **Silaidos C, Pilatus U, Grewal R, Matura S, Lienerth B, Pantel J, Eckert GP.** Sex-associated differences in mitochondrial function in human peripheral blood mononuclear cells (PBMCs) and brain. *Biol Sex Differ* 9: 34, 2018. doi:10.1186/s13293-018-0193-7.

68. **Silverman EK, Weiss ST, Drazen JM, Chapman HA, Carey V, Campbell EJ, Denish P, Silverman RA, Celedon JC, Reilly JJ, Ginns LC, Speizer FE.** Gender-related differences in severe, early-onset chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 162: 2152–2158, 2000. doi:10.1164/ajrccm.162.6.2003112.

69. **Svanes C, Omenaas E, Jarvis D, Chinn S, Gulsvik A, Burney P.** Parental smoking in childhood and adult obstructive lung disease: results from the European Community Respiratory Health Survey. *Thorax* 59: 295–302, 2004. doi:10.1136/thx.2003.009746.

70. **Tam A, Morrish D, Wadsworth S, Dorscheid D, Man SF, Sin DD.** The role of female hormones on lung function in chronic lung diseases. *BMC Womens Health* 11: 24, 2011. doi:10.1186/1472-6874-11-24.

71. **Tettamanti G, Ljung R, Mathiesen T, Schwartzbaum J, Feychting M.** Maternal smoking during pregnancy and the risk of childhood brain tumors: Results from a Swedish cohort study. *Cancer Epidemiol* 40: 67–72, 2016. doi:10.1016/j.canep.2015.11.009.

72. **Torén K, Hermansson B-A.** Incidence rate of adult-onset asthma in relation to age, sex, atopy and smoking: a Swedish population-based study of 15813 adults. *Int J Tuberc Lung Dis* 3: 192–197, 1999.

73. **Toriola AT, Vääräsmäki M, Lehtinen M, Zeleniuch-Jacquotte A, Lundin E, Rodgers K-G, Lakso H-A, Chen T, Schock H, Hallmans G, Pukkala E, Toniolo P, Grankvist K, Surcel HM, Lukanova A.** Determinants of maternal sex steroids during the first half of pregnancy. *Obstet Gynecol* 118: 1029–1036, 2011. doi:10.1097/AOG.0b013e3182342b7f.

74. **Townsend EA, Miller VM, Prakash YS.** Sex differences and sex steroids in lung health and disease. *Endocr Rev* 33: 1–47, 2012. doi:10.1210/er.2010-0031.

75. **Ventura-Clapier R, Moulin M, Piquereau J, Lemaire C, Mericskay M, Veksler V, Garnier A.** Mitochondria: a central target for sex differences in pathologies. *Clin Sci (Lond)* 131: 803–822, 2017. doi:10.1042/CS20160485.

76. **Vienneau D, Infanger D, Feychting M, Schüz J, Schmidt LS, Poulsen AH, Tettamanti G, Klæboe L, Kuehni CE, Tynes T, Von der Weid N, Lannering B, Röösli M.** A multinational case-control study on childhood brain tumours, anthropogenic factors, birth characteristics and prenatal exposures: a validation of interview data. *Cancer Epidemiol* 40: 52–59, 2016. doi:10.1016/j.canep.2015.11.006.

77. **Vrijlandt EJ, Gerritsen J, Boezen HM, Duiverman EJ; Dutch POPS-19 Collaborative Study Group.** Gender differences in respiratory symptoms in 19-year-old adults born preterm. *Respir Res* 6: 117, 2005. doi:10.1186/1465-9921-6-117.

78. **Wagner NJ, Camerota M, Propper C.** Prevalence and perceptions of electronic cigarette use during pregnancy. *Matern Child Health J* 21: 1655–1661, 2017. doi:10.1007/s10995-016-2257-9.

79. **Xiao D, Xu Z, Huang X, Longo LD, Yang S, Zhang L.** Prenatal gender-related nicotine exposure increases blood pressure response to angiotensin II in adult offspring. *Hypertension* 51: 1239–1247, 2008. doi:10.1161/HYPERTENSIONAHA.107.106203.

80. **Young S, Sherrill DL, Arnott J, Diepeveen D, LeSouëf PN, Landau LI.** Parental factors affecting respiratory function during the first year of life. *Pediatr Pulmonol* 29: 331–340, 2000. doi:10.1002/(SICI)1099-0496(200005)29:5< a331:AID-PPUL1> a3.0.CO;2-A.

81. **Zakarya R, Adcock I, Oliver BG.** Epigenetic impacts of maternal tobacco and e-vapour exposure on the offspring lung. *Clin Epigenetics* 11: 32, 2019. doi:10.1186/s13148-019-0631-3.

82. **Zeilinger S, Kühnel B, Klopp N, Baurecht H, Kleinschmidt A, Gieger C, Weidinger S, Lattka E, Adamski J, Peters A, Strauch K, Waldenberger M, Illig T.** Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One* 8: e63812, 2013. doi:10.1371/journal.pone.0063812.

28

1.3 Outline of Chapters

A review of the literature **(Chapter 1)** demonstrated that sexual dimorphism is prevalent across multiple respiratory diseases, particularly asthma and chronic obstructive pulmonary disease (COPD). **Chapter 1** consists of two peer-reviewed and accepted review articles exploring different aspects of sex differences in respiratory diseases. The review titled 'Sexual dimorphism in chronic respiratory diseases' **(Section 1.1)** delineates epidemiological, cellular and molecular patterns of sex differences in asthma, COPD and lung cancer. The other review, titled 'Sex-specific effects of *in utero* and adult tobacco smoke exposure' **(Section 1.2)**, highlights the complex and dynamic effect of cigarette smoking on the development of various diseases between the sexes. These reviews recognise that despite clear epidemiological evidence, the genetic and molecular mechanisms driving sex differences remain poorly understood and underinvestigated. Multiple factors have been posited to contribute to differences between males and females, such as physiological differences in lung size, sex hormone levels and sex chromosomes. A range of studies demonstrate that these factors affect fundamental processes in asthma and COPD – aberrant inflammation, altered fibrosis and airway remodelling and dysregulation of cell death. At the beginning of this project, it was necessary first to identify whether sex differences in these characteristic features of the disease are observable at a cellular level. If cells from males and females demonstrate dissimilar regulation of pathological processes, this may contribute to reported epidemiological differences in disease susceptibility and severity.

As such, we investigated whether sex differences in regulating the proinflammatory response existed *ex vivo*, irrespective of patient disease diagnosis. This study, included in **Chapter 2**, demonstrated that distinct immunoregulatory differences exist between cells isolated from males and females and are pathway-specific. As these cells have been removed from the human biological environment for a significant time, they have also been removed from direct exposure to circulating sex hormones (estrogen and testosterone). Therefore, we concluded an intrinsic factor that is different between males and females is driving the observed differences. This realisation drew our attention towards the sex chromosomes for reasons elucidated in the literature reviews **(Chapter 1)**.

Chapter 2 showed that cells isolated from males and females have different proinflammatory regulation mechanisms, irrespective of disease diagnosis. These cells are removed from the direct effect of circulating biological factors, indicating that factors inherently different between cell from males and females may cause differences disease-relevant processes. This notion pointed us towards the sex chromosomes. Genes on the sex chromosomes are often overlooked due to the complex processes and phenomena that compensate for females carrying two X-chromosomes and males carrying one X and one Y chromosome. Pivotal

publications by Bellott et al. [1], Grath and Parsch [2], and Bradbury [3] highlighted potential genome regulatory functions of *ZFX, ZFY, RPS4Y1, RPS4X, UTX* and *UTY.* These genes are all ubiquitously expressed, indicating evolutionary and biological importance. Furthermore, they regulate critical processes such as gene transcription, protein translation and changes to the epigenome. Importantly, these genes exist as pairs, with similar but non-exact copies on the X chromosome (i.e. *ZFX, RPS4X* and *UTX*) and Y chromosome (i.e. *ZFY, RPS4Y1* and *UTY*). Generally, they are assumed to have equal functions, but this is yet to be confidently proven in the literature. The functions of these genes have not been characterised completely, nor have they been widely compared for differential functions. The current known features and functions of these gene pairs is described and discussed in detail within the introduction of the relevant experimental chapters.

As such, we endeavoured to explore and compare the function of these candidate genes located on the X and Y chromosome (*ZFX, ZFY, RPS4Y1, UTX* and *UTY*) using CRISPR-Cas9-generated knockout cell lines. A wide array of techniques were required to complete the required functional studies and gene expression analyses. **Chapter 3** contains a detailed description of the methodology and analysis techniques used throughout this thesis and will be referenced throughout the results chapters. An important facet of this project is the bioinformatic investigation of RNA-sequencing and proteomics analyses. This skill was developed by analysis of an external dataset, resulting in a publication included as **Chapter 4**. This manuscript functioned as a tool to develop my bioinformatic toolset and to advance my written communication skills. Thus, it has been included in this thesis as it was critical to completing this work.

Understanding the similarities and dissimilarities in the function of these genes between males and females makes it possible to uncover novel pathways driving sex differences and identify clinical targets in asthma and COPD. **Chapter 5** identifies novel functions and highlights an important role for ZFX and ZFY in asthma. **Chapter 6** explores the link between RPS4Y1 and asthma. Finally, **Chapter 7** highlights the complex interaction and function of UTX and UTY in COPD.

1.4 Hypothesis & Aims

Asthma and COPD are prominent respiratory diseases that place a significant burden on society and healthcare systems worldwide. Both diseases demonstrate clear epidemiological and pathophysiological patterns of sexual dimorphism. The genetic and molecular factors contributing to these sex differences remain unknown. By understanding the fundamental mechanisms causing distinct disease phenotypes between males and females, the aetiology and process driving asthma and COPD may begin to be uncovered. The sex hormones estrogen and testosterone are key biological factors mediating sexually dimorphic responses and patterns in disease. However, the mechanisms by which estrogen and testosterone regulate disease processes remain incompletely elucidated [4]. The sex chromosome complement between males and females presents a distinct imbalance of gene dosage. Important genome regulators exist on the X and Y chromosome, potentially contributing to sex-biased gene expression [1, 2]. The molecular and physiological consequences of the gene imbalance between the sexes remains poorly understood and understudied.

We hypothesise that homologous gene pairs located on the X and Y chromosomes regulate hallmark disease features of inflammation, fibrosis and cell death. Furthermore, these gene pairs will not have equivalent expression levels or function. Thus, imbalanced expression and function of these genes in regulating key pathological processes may contribute to differences in the pathogenesis and progression of asthma and COPD between males and females.

This thesis aims to define the role of X and Y chromosome-linked gene pairs (*ZFX*/*ZFY*, *RPS4Y1*/*RPS4X* and *UTX/UTY*) in regulating inflammation, fibrosis and remodelling and cell death processes.

The specific aims to investigate this overall aim are:

1) Establish whether disease-relevant differences can be identified between male and female cells when removed from circulating biological factors
2) Define the function of each XY gene pair in regulating inflammatory, fibrotic, remodelling and cell death mechanisms
3) Compare and determine whether the expression and function of XY gene pairs is equivalent
4) Assess and explore the pathways mediated by XY gene pairs in the context of asthma or COPD

These aims will be explored for each gene pair throughout **Chapters 5, 6 and 7**.

Chapter 2

2.1 Publication Declaration - 'Sexually dimorphic production of interleukin-6 in respiratory disease'

Karosham D. Reddy, Sandra Rutting, Katrina Tonga, Dikaia Xenaki, Jodie L. Simpson, Vanessa M. McDonald, Marshall Plit, Monique Malouf, Razia Zakarya & Brian G.G. Oliver (2020), *Physiological Reports*.

Status: Published.

Author Contributions: KDR, BGO, and RZ conceived the idea. KDR, SR, KT, DX, MP, and MM contributed to data acquisition. KDR, BGO, JS, and VM performed, verified and discussed data analysis and interpretation. All authors discussed and contributed to the drafted manuscript for intellectual content.

Signatures:

Name:	Signature:	Date:
Karosham D. Reddy	Production Note: Signature removed prior to publication.	**08/12/2022**
Sandra Rutting	Production Note: Signature removed prior to publication.	**16/01/2023**
Katrina Tonga	Production Note: Signature removed prior to publication.	**12/12/2022**
Dikaia Xenaki	Production Note: Signature removed prior to publication.	**09/12/2022**
Jodie L. Simpson	Production Note: Signature removed prior to publication.	**08/12/2022**
Vanessa M. McDonald	Production Note: Signature removed prior to publication.	**10/01/2023**
Marshall Plit	Production Note: Signature removed prior to publication.	**14/12/2022**
Monique Malouf	Production Note: Signature removed prior to publication.	**14/12/2022**
Razia Zakarya	Production Note: Signature removed prior to publication.	**08/12/2022**
Brian G.G. Oliver	Production Note: Signature removed prior to publication.	**08/12/2022**

ORIGINAL RESEARCH

Sexually dimorphic production of interleukin-6 in respiratory disease

Karosham D. Reddy[1,2] [ID] | Sandra Rutting[2] | Katrina Tonga[2,3] | Dikaia Xenaki[2] | Jodie L. Simpson[4] | Vanessa M. McDonald[4] | Marshall Plit[3] | Monique Malouf[3] | Razia Zakarya[1,2] | Brian G. Oliver[1,2]

[1]School of Life Sciences, University of Technology Sydney, Sydney, NSW, Australia

[2]Respiratory Cellular and Molecular Biology, Woolcock Institute of Medical Research, University of Sydney, Sydney, NSW, Australia

[3]St Vincent's Hospital Sydney and St Vincent's Clinical School, University of New South Wales, Darlinghurst, NSW, Australia

[4]Priority Research Centre for Healthy Lungs, Faculty of Health and Medicine, University of Newcastle, Callaghan, NSW, Australia

Correspondence
Karosham D. Reddy, Woolcock Institute of Medical Research, Cells Group, Level 3, 431 Glebe Point Road, Glebe, NSW 2037, Australia.
Email: karosham.reddy@woolcock.org.au

Funding information
This study was supported by funding from the National Health and Medical Research Council (NHMRC). KDR is supported by a postgraduate research fellowship from the University of Technology Sydney.

Abstract

Diverging susceptibility and severity in respiratory diseases is prevalent between males and females. Sex hormones have inconclusively been attributed as the cause of these differences, however, strong evidence exists promoting genetic factors leading to sexual dimorphism. As such, we investigate differential proinflammatory cytokine (interleukin (IL)-6 and CXCL8) release from TNF-α stimulated primary human lung fibroblasts in vitro. We present, for the first time, in vitro evidence supporting clinical findings of differential production of IL-6 between males and females across various respiratory diseases. IL-6 was found to be produced approximately two times more from fibroblasts derived from females compared to males. As such we demonstrate sexual dimorphism in cytokine production of IL-6 outside the context of biological factors in the human body. As such, our data highlight that differences exist between males and females in the absence of sex hormones. We, for the first time, demonstrate inherent in vitro differences exist between males and females in pulmonary fibroblasts.

KEYWORDS
fibroblasts, inflammation, respiratory disease, sex-specific response

1 | INTRODUCTION

Sexual dimorphism occurs in various pathologies such as cardiovascular disease, cancer, and respiratory conditions (Alexander, Dasinger, & Intapad, 2011; Lopes-Ramos et al., 2018; Raghavan & Jain, 2016); influencing pathogenesis, progression, and response to treatment. This phenomenon is prominent in respiratory illnesses such as chronic obstructive pulmonary disease (COPD) and asthma. There is evidence that females have a preponderance for developing COPD, but the mechanism is not understood (Eisner et al., 2005; Tam et al., 2016). Furthermore, prepubescent males have higher prevalence of asthma (Almqvist, Worm, Leynaert, & 'GENDER', W. G. O. G. L. W., 2008).

33

This susceptibility is attributed to male airway growth lagging that of the parenchyma (dysanapsis) restricting expiratory rate, whereas females have accelerated lung growth and increased airway size (Almqvist et al., 2008). Therefore, clear differences exist between the sexes in the prevalence of respiratory disease, however, the exact mechanism remains unknown.

One of the best-described examples of sex difference in a disease is in asthma where a shift occurs during puberty; female asthma incidence increases over-and-above males (Almqvist et al., 2008). Age-associated changes in sex-hormone levels are often attributed as the cause. In fact, estrogen and testosterone have opposing effects on the immune system; immunocompetence and suppression, respectively (Fish, 2008). Similarly, sex hormones have been ascribed to cause differences between male and female COPD phenotypes; affecting oxidative stress pathways and airway remodeling (Tam et al., 2016).

However, sex hormones cannot fully explain sex differences in respiratory disease. Sexual dimorphism occurs before gonadal development; well before sex-hormone production (Deegan & Engel, 2019; Werner et al., 2017). Males and females demonstrate distinct responses to their environment at this early stage, reflecting clear sexual identity informed by genetic factors (Deegan & Engel, 2019). Recent GWAS evidence suggests X-chromosome miRNA contributes to asthma onset at different stages of development (Ferreira et al., 2019). Therefore, a hormone-independent mechanism can drive sex differences. It remains unclear whether nonhormonal factors significantly contribute to disease later in life.

Tumor necrosis factor-alpha (TNF-α) is recognized as a potent proinflammatory cytokine in various respiratory diseases. TNF-α is well-known to stimulate production of interleukin (IL)-6 and C-X-C motif ligand 8 (CXCL8), prominent chemo-attractant cytokines which induce neutrophil infiltration and activation, driving inflammation in the lungs (Lundblad et al., 2005). This study aimed to investigate if IL-6 and CXCL8 production differ in pulmonary fibroblasts derived from male and female patients in-vitro when stimulated by TNF-α.

2 | MATERIALS AND METHODS

2.1 | Patients

Primary fibroblasts were isolated from 36 samples of lung parenchyma from patients with a variety of diagnoses. Each diagnosis was made by thoracic physicians according to current guidelines. Protocols were submitted to and approved by a human research ethics committee and prior written and informed consent was obtained from patients under approval

TABLE 1 Summary patient demographics

	Male	Female
n	21	15
Mean age ($\pm SD$)	58.3 ($\pm$12.1)	53.2 ($\pm$15.3)
Mean FEV_1/FVC ($\pm SD$)	0.61 ($\pm$0.28)	0.48 ($\pm$0.21)
Smokers/nonsmokers/ unknown	16/3/2	8/3/4
Pathology		
COPD (GOLD stage 4)	$n = 6$	$n = 6$
Idiopathic pulmonary fibrosis	$n = 8$	–
Thoracic malignancy	$n = 4$	$n = 3$
Bronchiolitis	$n = 1$	$n = 3$
Pulmonary hypertension	$n = 1$	–
Bronchiectasis	–	$n = 1$
Eisenmenger syndrome	–	$n = 1$
Pneumonitis	–	$n = 1$
No diagnosis	$n = 1$	–

by code #X14-0045. Patient demographics are summarized in Table 1.

2.2 | Cell culture

Primary lung fibroblasts were isolated from human lung tissue, as previously described by (Krimmer, Ichimaru, Burgess, Black, & Oliver, 2013). Cells were grown in vitro, seeded at a density of 6.2×10^{-4} cells/ml in 12-well plates in Dulbecco's Modified Eagles Medium (DMEM) (Gibco) containing 5% fetal bovine serum (FBS), 25 mM Hepes buffer (Gibco), and 1% antibiotic-antimycotic (Gibco) at 37°C/5% CO_2. Once cells reached 80% confluency, they were serum starved in DMEM supplemented with 0.1% bovine serum albumin (BSA) (Sigma-Aldrich), 25 mM Hepes buffer, and 1% antibiotic-antimycotic for 24 hr prior to stimulation. Fibroblasts cultures between passages 2 and 4. The use of early passages attempts to ensure that cell health and processes are maintained as much as possible. Mycoplasma testing was completed on all cell-lines and returned a negative result.

2.3 | Cell Stimulation with TNF-α

Isolated primary fibroblasts were stimulated with TNF-α (1 ng/ml) (ThermoFisher #T0157) or vehicle control (0.1% BSA) for 24 hr. All cells were incubated at 37°C/5% CO_2 for 48 hr. Cell-free supernatants were collected and stored at −20°C until further analysis. IL-6 and IL-8 production were measured in cell-free supernatant by ELISA.

2.4 | Measurement of IL-6 and CXCL8 levels

Sandwich ELISA was used on cell-free supernatants to measure the level of IL-6 and CXCL8 cytokines as described by (Rutting et al., 2018).

2.5 | Statistical analysis

Statistical analysis was completed using GraphPad Prism version 8 software (GraphPad Software). Comparisons were carried out on the data by Student's parametric two-tailed t-test. All data on figures are presented as mean $\pm$ standard error of the mean (SEM). Statistical significance was determined at $p < .05$.

3 | RESULTS

3.1 | Fibroblasts from male and female donors demonstrate different responses To TNF-A stimulation

IL-6 and CXCL8 production was measured in cell-free supernatant by ELISA. No difference in baseline production of either cytokine was seen by pulmonary fibroblasts between male and female patients; IL-6:133.0 $\pm$ 17.06 pg/ml versus 98.80 $\pm$ 21.78 pg/ml and CXCL8: 36.57 $\pm$ 4.22 pg/ml versus 33.74 $\pm$ 5.32 pg/ml, respectively (Figure 1a and b). Similarly, no significant difference was observed between male and female derived fibroblast when stimulated by TNF-α. However, female derived fibroblasts produced a greater fold-change from baseline increase in IL-6 production than males; 95.15 $\pm$ 17.27 versus 53.94 $\pm$ 29.94 ($p = .016$), respectively (Figure 1c). This effect was observable irrespective of disease. Conversely, no difference was observed between the sexes when CXCL8 was investigated in the same manner (Figure 1d). These results were reflected in the subpopulation of COPD only diagnoses, where fibroblasts derived from female patients showed greater fold-change in IL-6 production compared to male derived cells; 98.11 $\pm$ 11.70 versus 46.49 $\pm$ 4.38 ($p = .002$), respectively (Figure 1e). Regardless of respiratory diagnosis, females produce almost double the IL-6 production due to TNF-α stimulation. Upon removal of the fibroblasts from males with an IPF diagnosis, the trend toward greater fold-IL-6 production from females was maintained, although not significant; $p = .061$ (data not shown). No differences in TNF-α induced CXCL8 production were observed suggesting cytokine-specific sexual dimorphism can occur.

4 | DISCUSSION

We show a sexually dimorphic expression pattern exists in fibroblasts removed from the human body. These cells were removed from natural biological influences of hormones and grown for 2–3 months in vitro, indicating an intrinsic mechanism contributes to the sexually dimorphic production of IL-6 in respiratory disease. Female cells were more liable toward a greater induction of cytokine production compared to males, which may be associated with their generally worse prognosis in respiratory disease. Therefore, our data suggest a hormone-independent regulatory mechanism exists between the sexes. This is the first study to demonstrate this phenomenon in vitro in primary human pulmonary fibroblasts.

A similar phenomenon has been observed in both murine and human in-vivo studies. When exposed to ozone, female mice demonstrated increased *IL-6* expression, among other inflammatory genes, compared to males (Cabello et al., 2015). Although, this difference occurs outside of a disease state. Furthermore, IL-6 among other cytokines are reported to differ between the sexes in COPD (de Torres et al., 2011). This, when compared with our results supports the potential of disease affecting male and female immune responses differently. These studies highlight sexually dimorphic gene regulation in response to stimuli which may contribute to respiratory disease processes. Hence, these studies in conjunction with our findings suggest sex differences may be driven by an internal cellular mechanism.

Our data support studies which indicate tissue-specific sexually dimorphic regulation. This is well-described for the gonads; the primary tissue where sex differentiation exists. This bias is evolutionarily conserved through multiple species and taxa from flies to primates, recognized at the mRNA level (Ober, Loisel, & Gilad, 2008). Importantly, sexually dimorphic expression is reported for autosomal genes, indicating a complex regulatory network is contributing to sex differences. Importantly, the sexually dimorphic expression has also been reported for genes encoded on autosomes in a tissue-specific manner. The evolution of these genes is asserted to be driven by sex-specific pressures, increasing sex-bias over time (Reinius et al., 2008). Therefore, the same exposure would result in different gene sets being utilized between the sexes.

It is important to acknowledge that these differences may be a consequence of the influence of epigenetic marks induced by sex hormones which have been maintained in vitro (Nugent et al., 2015). Fibroblasts derived from various locations in the body have been characterized to express sex steroid hormones receptors, including estrogen receptor alpha (ERα) (Mukudai et al., 2015). Further, a dynamic interplay between hormones and epigenetic patterns has been established. In fact, estrogen is described to exert an epigenetic influence on gene expression (Asai et al., 2001;
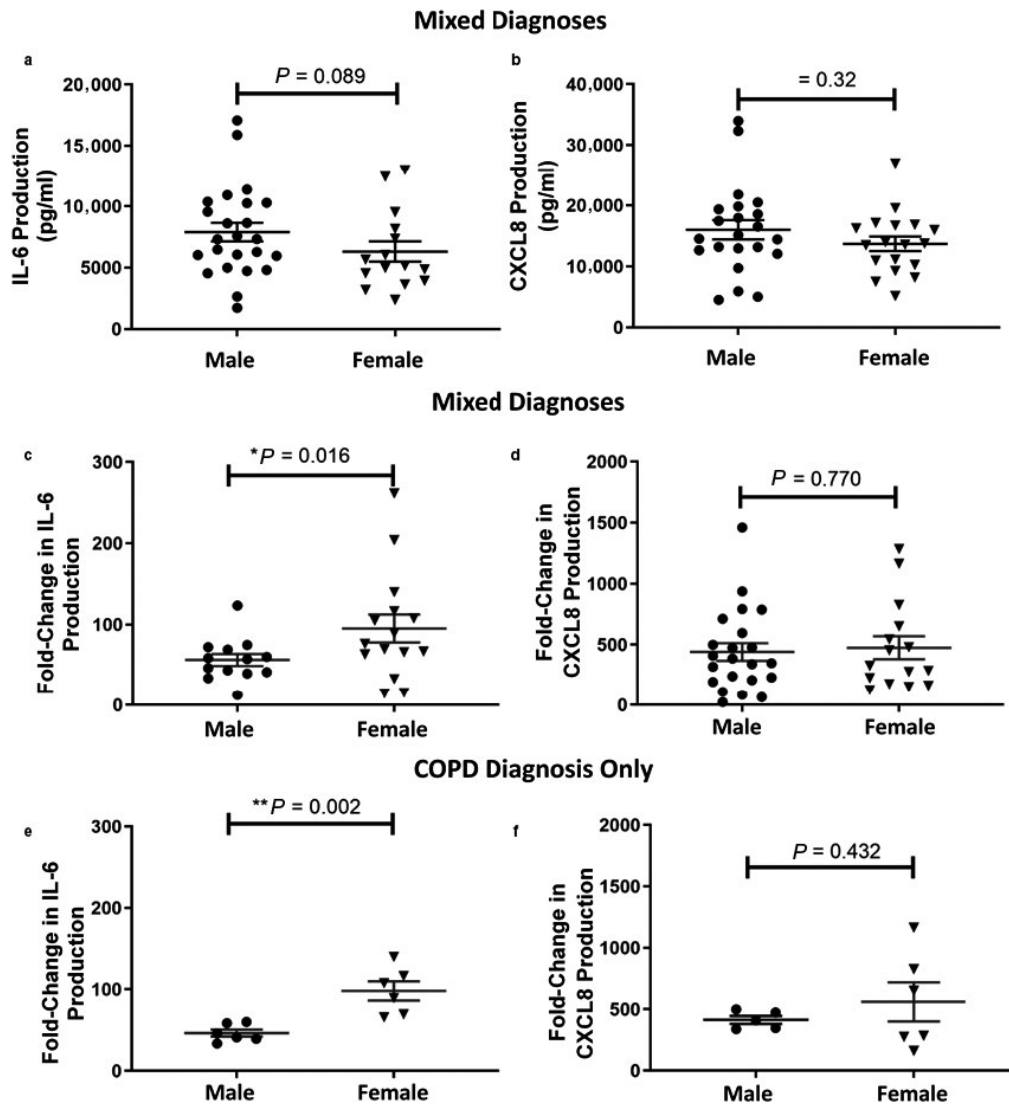
FIGURE 1 Effect of TNF-α on IL-6 and CXCL8 production from pulmonary fibroblasts. Cells were treated with TNF-α (1 ng/ml) for 24 hr. TNF-α induced IL-6 and CXCL8 production (a and b). Fold-change determined by comparison of TNF-α stimulated to untreated cells (c and d). Fold-change determined by comparison of TNF-α stimulated to untreated cells from patients with COPD (e and f). Data are represented as mean ± SEM and analyzed using a Student's parametric two-tailed t-test; $n = 5–21$. Statistical significance is indicated by $*p < .05$ and $**p < .01$

Zhang & Ho, 2011). TNF-α signaling is influenced by estrogen, subsequently impacting the immune response (Song, Kim, Kim, Lee, & Surh, 2019). Thus, epigenetic patterns imposed by sex hormones during the patient's lives could be maintained ex vivo, and influence the observed sexual dimorphic pattern reported in this study. However, the locality and longevity of such epigenetic marks is yet to be understood. Most studies looking at this interplay focus on the brain with limited work in the lung. It is prudent for future investigation to focus on the hormone-epigenetic interplay, as this will offer insight into the complex genetic-epigenetic mechanisms in disease.

Our study has limitations. Smoking history was only available for 83% of patients, the majority were ex-smokers ($n = 24$), and few were never-smokers ($n = 6$). However, the population diagnoses include smoking and nonsmoking-related diseases, reducing the likelihood of smoking functioning as a determining factor. Nonetheless, the potential contribution of smoking requires further investigation. Females generate an increased immune response to tobacco (Kynyk, Mastronarde, & McCallister, 2011). However, the driving mechanism remains unknown. Tobacco alters the epigenome in a sex-specific manner, presenting a broader mechanism of action (Ladd-Acosta et al., 2016). The X-chromosome contains the largest set of immune-related genes, with those that escape X-inactivation possibly contributing to this phenomenon. As such, the regulatory genome is sexually dimorphic (Ober et al., 2008), necessitating careful investigation to determine the mechanism for sex differences in gene regulation.

5 | CONCLUSIONS

Here, we present for the first time sexually dimorphic IL-6 production in-vitro. We speculate this difference is driven by either conserved genetic predisposition or epigenetic regulation of transcription. However, it is possible this effect is due to a continued hormone imprint on the genome; therefore, a detailed investigation is required. Our study shows that differential regulatory mechanisms exist between the sexes and is maintained outside of the body. As such, we highlight the importance of reporting sexual dimorphism in all investigations.

CONFLICT OF INTEREST
The authors have nothing to disclose.

AUTHORS' CONTRIBUTIONS
KDR, BGO, and RZ conceived the idea. KDR, SR, KT, DX, MP, and MM contributed to data acquisition. KDR, BGO, JS, and VM performed, verified and discussed data analysis and interpretation. All authors discussed and contributed to the drafted manuscript for intellectual content.

ORCID
Karosham D. Reddy https://orcid.org/0000-0001-9002-6930

REFERENCES

Alexander, B. T., Dasinger, J. H., & Intapad, S. (2011). Fetal programming and cardiovascular pathology. *Comprehensive Physiology*, *5*, 997–1025.

Almqvist, C., Worm, M., Leynaert, B, & GENDER', W. G. O. G. L. W. (2008). Impact of gender on asthma in childhood and adolescence: A GA2LEN review. *Allergy*, *63*, 47–57.

Asai, K., Hiki, N., Mimura, Y., Ogawa, T., Unou, K., & Kaminishi, M. (2001). Gender differences in cytokine secretion by human peripheral blood mononuclear cells: Role of estrogen in modulating lps-induced cytokine secretion in an ex vivo septic model. *Shock*, *16*(5), 340–343. https://doi.org/10.1097/00024382-20011 6050-00003

Cabello, N., Mishra, V., Sinha, U., Diangelo, S. L., Chroneos, Z. C., Ekpa, N. A., … Silveyra, P. (2015). Sex differences in the expression of lung inflammatory mediators in response to ozone. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, *309*, L1150–L1163.

de Torres, J. P., Casanova, C., Pinto-Plata, V., Varo, N., Restituto, P., Cordoba-Lanus, E., … Celli, B. R. (2011). Gender differences in plasma biomarker levels in a cohort of COPD patients: A pilot study. *PLoS ONE*, *6*, e16021. https://doi.org/10.1371/journ al.pone.0016021

Deegan, D. F., & Engel, N. (2019). Sexual dimorphism in the age of genomics: How, when, where. *Frontiers in Cell and Developmental Biology*, *7*, 186.

Eisner, M. D., Balmes, J., Katz, P. P., Trupin, L., Yelin, E. H., & Blanc, P. D. (2005). Lifetime environmental tobacco smoke exposure and the risk of chronic obstructive pulmonary disease. *Environmental Health*, *4*, 7.

Ferreira, M. A., Mathur, R., Vonk, J. M., Szwajda, A., Brumpton, B., Granell, R., … Jiang, Y. (2019). Genetic architectures of childhood-and adult-onset asthma are partly distinct. *The American Journal of Human Genetics*, *104*, 665–684.

Fish, E. N. (2008). The X-files in immunity: Sex-based differences predispose immune responses. *Nature Reviews Immunology*, *8*, 737.

Krimmer, D., Ichimaru, Y., Burgess, J., Black, J., & Oliver, B. (2013). Exposure to biomass smoke extract enhances fibronectin release from fibroblasts. *PLoS ONE*, *8*(12), 1–10.

Kynyk, J. A., Mastronarde, J. G., & McCallister, J. W. (2011). Asthma, the sex difference. *Current Opinion in Pulmonary Medicine*, *17*, 6–11.

Ladd-Acosta, C., Shu, C., Lee, B. K., Gidaya, N., Singer, A., Schieve, L. A., … Windham, G. C. (2016). Presence of an epigenetic signature of prenatal cigarette smoke exposure in childhood. *Environmental Research*, *144*, 139–148.

Lopes-Ramos, C. M., Kuijjer, M. L., Ogino, S., Fuchs, C. S., Demeo, D. L., Glass, K., & Quackenbush, J. (2018). Gene regulatory network analysis identifies sex-linked differences in colon cancer drug metabolism. *Cancer Research*, *78*, 5538–5547.

Lundblad, L. K., Thompson-Figueroa, J., Leclair, T., Sullivan, M. J., Poynter, M. E., Irvin, C. G., & Bates, J. H. (2005). Tumor necrosis factor–α overexpression in lung disease: A single cause behind a complex phenotype. *American Journal of Respiratory and Critical Care Medicine*, *171*, 1363–1370.

Mukudai, S., Matsuda, K. I., Nishio, T., Sugiyama, Y., Bando, H., Hirota, R., … Kawata, M. (2015). Differential responses to steroid hormones in fibroblasts from the vocal fold, trachea, and esophagus. *Endocrinology*, *156*, 1000–1009.

Physiological Reports

Nugent, B. M., Wright, C. L., Shetty, A. C., Hodes, G. E., Lenz, K. M., Mahurkar, A., ... McCarthy, M. M. (2015). Brain feminization requires active repression of masculinization via DNA methylation. *Nature Neuroscience*, *18*, 690.

Ober, C., Loisel, D. A., & Gilad, Y. (2008). Sex-specific genetic architecture of human disease. *Nature Reviews Genetics*, *9*, 911.

Raghavan, D., & Jain, R. (2016). Increasing awareness of sex differences in airway diseases. *Respirology*, *21*, 449–459.

Reinius, B., Saetre, P., Leonard, J. A., Blekhman, R., Merino-Martinez, R., Gilad, Y., & Jazin, E. (2008). An evolutionarily conserved sexual signature in the primate brain. *PLoS Genetics*, *4*, e1000100.

Rutting, S., Papanicolaou, M., Xenaki, D., Wood, L. G., Mullin, A. M., Hansbro, P. M., & Oliver, B. G. (2018). Dietary ω-6 polyunsaturated fatty acid arachidonic acid increases inflammation, but inhibits ECM protein expression in COPD. *Respiratory Research*, *19*, 211.

Song, C.-H., Kim, N., Kim, D.-H., Lee, H.-N., & Surh, Y.-J. (2019). 17-β estradiol exerts anti-inflammatory effects through activation of Nrf2 in mouse embryonic fibroblasts. *PLoS ONE*, *14*(8), 1–13.

Tam, A., Churg, A., Wright, J. L., Zhou, S., Kirby, M., Coxson, H. O., ... Sin, D. D. (2016). Sex differences in airway remodeling in a mouse model of chronic obstructive pulmonary disease. *American Journal of Respiratory and Critical Care Medicine*, *193*, 825–834.

Werner, R. J., Schultz, B. M., Huhn, J. M., Jelinek, J., Madzo, J., & Engel, N. (2017). Sex chromosomes drive gene expression and regulatory dimorphisms in mouse embryonic stem cells. *Biology of Sex Differences*, *8*, 28.

Zhang, X., & Ho, S.-M. (2011). Epigenetics meets endocrinology. *Journal of Molecular Endocrinology*, *46*, R11.

Chapter 3 – Methodology

This chapter will outline the methodologies used to generate the results presented in chapters 5, 6 and 7.

3.1 Cell culture

A549 (adenocarcinoma human alveolar basal epithelial cells) from ATCC (American Type Culture Collection) were grown in T75 tissue culture flasks. Cells were maintained in "growth medium"; Dulbecco's modified eagle medium (DMEM) (#31600091, ThermoFisher Scientific, MA, USA) supplemented with 10% fetal bovine serum (FBS, #16000044, ThermoFisher, MA, USA), 1% antibiotic/antimycotic (#15240096, ThermoFisher Scientific, MA, USA) and buffered with 25 mM HEPES (#BIOHB0265, Astral Scientific, NSW, AUS). This medium will be referred to as a 'growth medium' throughout the rest of this thesis. All incubation steps involving live cells are completed at 37°C/5% CO_2 throughout this project unless otherwise stated. Cells were regularly monitored by brightfield microscopy to ensure healthy and regular cell growth.

Cells were seeded at varying densities depending on the analysis or assay conducted, which will be specified as needed. Briefly, an 80% confluent T75 flask was washed with Hanks balanced salts solution (HBSS, #H2387, Sigma-Aldrich, Castle Hill, NSW) and incubated with 3 mL 0.05% (w/v) Trypsin (#15400054, LifeTechnologies, Waltham MA, USA) for 3 min at 37°C/5% CO_2. The trypsin was then inactivated with 7 mL of growth medium, with the cell suspension centrifuged at 1000 rpm for 5 min. The supernatant was aspirated, and cells were resuspended in growth medium. An aliquot of the suspension was removed to be stained with trypan blue (#15250061, Gibco, ThermoFisher, MA, USA) to identify dead cells. The aliquot was then loaded into a hemocytometer for manual cell counting. The necessary volume of suspension with the correct concentration of cells was calculated. Cells were seeded at the appropriate density in growth medium and incubated at 37°C/ 5% CO_2 until 80% confluency was reached in the cell culture plate.

3.2 Generating CRISPR Cas-9 knockout cell lines

3.2.1 Creation of CRISPR-Cas9 plasmid constructs

Guide RNA (gRNA) sequences were constructed using the online free design tool Benchling (www.benchling.com); this tool has been validated and used previously[5]. gRNA sequences for each gene of interest (*KDM6A*, *KDM6C*, *RPS4X*, *RPS4Y1*, *ZFX* and *ZFY*) were designed to target common exons present across all known splice variants as listed on www.ensembl.org. All gRNA sequences contained the protospacer adjacent motif (PAM) sequence "NGG". The ideal gRNA sequences were determined using the on-target and off-target scores generated by Benchling. PCR and sequencing primers were designed using the NCBI Primer-Blast online tool[6]. PCR primers were designed to produce a product of at least 800 base pairs in size spanning both exon and intronic regions, with the gRNA sequence located approximately in the middle. Forward sequencing primers were designed to bind to the DNA sequence at least 100 base pairs from the start of the PCR forward primer binding site. Table 3.1 contains the primers for each gene of interest ordered from Integrated DNA Technologies (IDT, IA, USA).

The pX458 (#48138, Addgene, MA, USA) green fluorescent protein (GFP) tagged plasmid was used as a vector for the guide RNA (gRNA). gRNA primers were annealed using T4 ligation buffer (#B0202S) and T4 polynucleotide ligase (#M0201S) (NEB, Ipswich, MA, USA) with temperature cycling through 37°C for 30 min, 95°C for 5 min before ramping the temperature down to 25°C at 5°C/min. The plasmid was then digested using the BbsI restriction enzyme (#FD1014, ThermoFisher Scientific, Waltham MA, USA). The annealed primers were ligated to the cut sites using T7 DNA ligase (NEB, Ipswich, MA, USA). The reaction was cycled through 37°C for 5 min and 25°C for 5 min six times before being held at 4°C. Unwanted recombination products were removed using PlasmidSafe exonuclease (#E3101K, Biosearch Technologies, Hoddesdon, UK) with incubation at 37°C for 30 min. Samples were stored at -20°C for use within 30 days.

Table 3.1: List of primer sequences used to generate knockout cell lines. For = Forward primer sequence; Rev = Reverse primer sequence

Target	Primers	Sequence
KDM6A	Guide RNA	For: CACCGCAAACCAAGACCATATAAAA Rev: AAACTTTTATATGGTCTTGGTTTGC
	PCR	For: TTCAGGGCTGGAATACCTCT Rev: AAACAACAATGGGGCAAAGGC
	Sequencing	TCTGTGCCATCTGATCACTGT
KDM6C	Guide RNA	For: CACCGCCTTGGCTCGACAAAAGCTG Rev: AAACCAGCTTTTGTCGAGCCAAGGC
	PCR	For: AACACCTTCACTTTCTCACCT Rev: TTAGCAACTGGCAGTCCAAAAG
	Sequencing	TGACGGAATTGAGAACAAAAGGATT
RPS4X	Guide RNA	For: CACCGATCGGGGTAGCGGATGGTG Rev: AAACCACCATCCGCTACCCCGATC
	PCR	For: CTGGGAAACTGGAGTGGCAT Rev: CTAGGCAGGAAGAAATAATCTGC
	Sequencing	TCTGCAGCAGTTAGGGAACC
RPS4Y1	Guide RNA	For: CACCGTTGCCTCTTCCACTGTGATG Rev: AAACCATCACAGTGGAAGAGGCAAC
	PCR	For: TGGGGCCAGGACATTGATTG Rev: GAGGCAGACATCAGGCAGTC
	Sequencing	GCAGGGGTTGGTTGGTTTCA
ZFX	Guide RNA	For: CACCGTATGCCAGAACACGTCTTGA Rev: AAACTCAAGACGTGTTCTGGCATAC
	PCR	For: TGACTGCCTCAATTGCTTTGTT Rev: TGACAATCACTGTACCCCAAGA
	Sequencing	GGAGCTAATCTGTTTTCCCAGTA
ZFY	Guide RNA	For: CACCGTGCATCGATGGTCACTCCAG Rev: AAACCTGGAGTGACCATCGATGCAC
	PCR	For: TTTTCAGATCTTCAAGAACATTGGT Rev: ACCTATCAGAGAATATGAAAGCTCC
	Sequencing	TCAGTGGATGATGCTGGCAAA
pX458	U6 sequencing	ACCGAAATATATAGAA

3.2.2 Bacterial transformation and plasmid amplification

Competent DH5α *Escherichia coli* (#C2987H, NEB, MA, USA) were transformed with the plasmid constructs generated in section 3.2.1. Heat shock at 42°C for 30 sec in a water bath was used to transform the *E. coli*. Bacteria were incubated on ice for 2 min to recover. To assist the bacterial recovery process, 125 µL of lysogeny broth (LB) (#12780052, ThermoFisher Scientific, Waltham MA, USA) was added, and the bacteria were incubated for 1 hour at 37°C and 230 rpm. Post-incubation, the bacterial culture was spread plated on ampicillin (100 µg/mL) (#611770250, ThermoFisher Scientific, MA, USA) positive agar plate. 5 µL of culture was also incubated on a plate without ampicillin to confirm a positive selection of transformed bacterial cells. Plates were incubated at 37°C overnight. The next day, a single bacterial colony was isolated and expanded in 20 mL of LB containing ampicillin (100 µg/mL) overnight at 37°C and 230 rpm. Following clonal expansion, plasmid DNA was extracted using the Qiagen plasmid extraction kit (#27104, Qiagen, Hilden, DEU), according to the manufacturer's instructions. Purified plasmids were stored at -20°C for use within 30 days.

3.2.3 Plasmid construct sequencing

To confirm the correct and complete insertion of the gRNA primers into the pX458 plasmid, purified plasmids from section 3.2.2 were analysed by sanger sequencing. Plasmid concentration was determined using a Nanodrop. According to the submission instructions provided by the Australian Genome Research Facility (AGRF, Westmead, NSW, AUS), 600 ng of plasmid was supplied with the pX458 sequencing primer (Table 3.1). Dendrogram results from AGRF were analysed using Benchling, demonstrating the correct incorporation of the gRNA sequence into the pX458 plasmid (Figure 3.1).
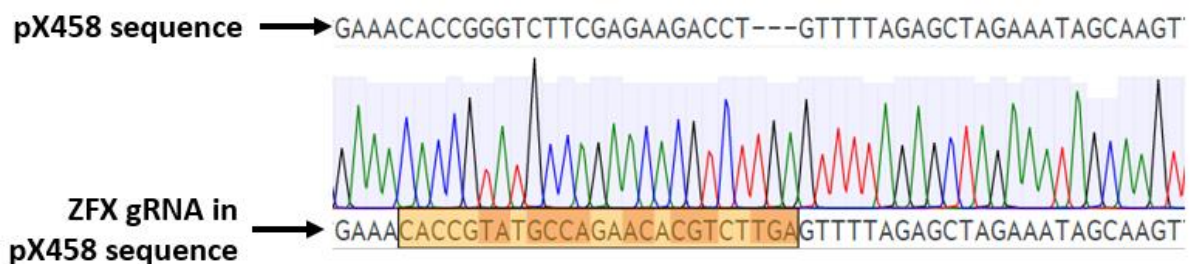
Figure 3.1: Example dendrogram generated by sanger sequencing confirming the incorporation of ZFX gRNA into the pX458 plasmid sequence. The yellow highlighted region is the ZFX guide RNA sequence integrated into the plasmid DNA.

3.2.4 Transfection of A549 cells

A549 cells were selected as they are a male-derived cell line, meaning they contain both X and Y chromosomes. This enabled an investigation of genes on both the X and Y chromosome. A549 cells were seeded in a 12-well cell culture plate (#150628, ThermoFisher Scientific, MA, USA) at 100,000 cells/well density in a DMEM growth medium. Cells were incubated at 37°C/ 5% CO_2 for 24 hours. Following this, cells were transfected with the plasmid construct using Lipofectamine 3000 reagent (#L3000001, ThermoFisher Scientific, MA, USA). The manufacturer's instructions were followed. In summary, one µg of plasmid construct was incubated with lipofectamine 3000 and p3000 reagents in Opti-MEM transfection medium (#11058021, ThermoFisher Scientific, MA, USA) for 15 min at room temperature. This allowed for the plasmids to be captured within the liposomes in the solution. The mixture was added on top of the cell culture with an equal volume of Opti-MEM medium. A control mixture consisting of only Opti-MEM was added to one well to evaluate whether any cell death was induced. A reagent control (RC) mixture of lipofectamine reagents without a plasmid construct was also included. This will be used as the wildtype cell line throughout future experiments. Cells were incubated for 5 hours at 37°C/5% CO_2 to allow for transfection to occur. After, 1 mL of growth media was added to assist the cell recovery. Cells were left at 37°C/5% CO_2 for a further 24 hours.

3.2.5 Single-cell sorting and selection for transfected cells

GFP expression was identified to be maximal at approximately 30 hours post-transfection (Figure 3.2). Therefore, this time point was selected to conduct single-cell sorting. Reagent control and transfected cells were washed with HBSS before being harvested using trypsin by incubation at 37°C/ 5% CO_2 for 5 min. After, phosphate-buffered saline solution supplemented with 10% FBS was added to the wells to inactivate the trypsin. The wells were washed with PBS-FBS solution to lift cells from the plate, with the cell suspension transferred to a sterile tube.
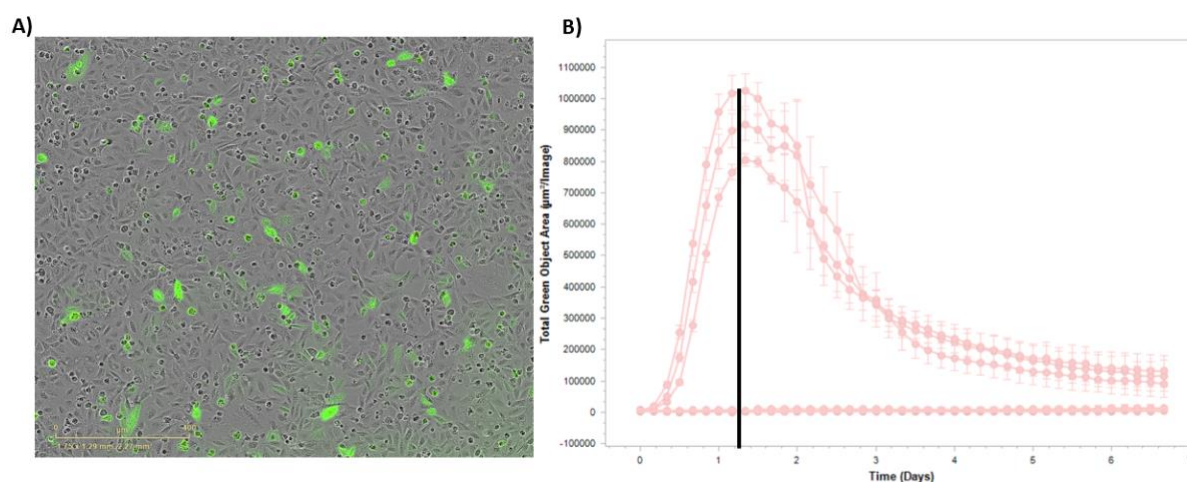
Figure 3.2: (A) Representative image of green fluorescent protein (GFP) expression from transfected cells. (B) The mean area of cells expressing GFP per image against time. The vertical black line indicates the time of maximal GFP expression from cells.

The BD FACS single-cell sorter (BD Biosciences, NJ, USA) was used to successfully identify transfected cells expressing the GFP protein. The manufacturer-provided instructions were followed to calibrate and set up the BD FACS sorter. Transfected A549 cells that expressed medium to high GFP signals were selected. The level of GFP expression was determined by comparison to the reagent control cell sample. A single A549 cell expressing GFP was identified and sorted into each well of a 96-well cell culture plate pre-filled with 100 µL of growth medium. The plate was then incubated at 37°C/ 5% CO_2 for one week to allow for clonal expansion of each cell per well. Once 70 – 80% confluency was reached, cells from the well were harvested and transferred to a 12-well plate before being moved to a T75 cell culture flask (#156499, ThermoFisher Scientific, MA, USA) once confluency was reached. DNA lysate was collected for downstream confirmation of knockout status (section 3.2.6). Cells were frozen and stored in liquid nitrogen until future use.

3.2.6 Sanger Sequencing to confirm knockout cell line

According to the manufacturer's instructions, the DNA lysates collected in 3.2.5 were purified using the Bioline Isolate II DNA extraction kit (#BIO-52066, Meridian Bioscience, OH, USA). DNA purity and amount were analysed using a Nanodrop. The region targeted by the gRNA sequence was amplified using the PCR primers as listed in Table 3.1, using the Hot Start *Taq* PCR kit (#M04962, NEB, MA, USA). The reactions were heated to 94°C for 2 min before being cycled through 94°C for 10sec, 59°C for 30 sec and 72°C for 2 min 35 times. The final extension of the amplification product was completed at 72°C for 7 min. PCR products were visualised on a 2% (w/v) agarose gel (#A9918, Sigma-Aldrich, Castle Hill, NSW) to confirm the correct region was amplified. The remaining post-PCR sample was purified and cleaned

using the Bioline PCR and Gel Purification Kit (#BIO-52059, Meridian Bioscience, OH, USA), before samples were analysed for quality and DNA quantity using a Nanodrop.

The amplified and clean PCR sequences were sent to AGRF for sanger sequencing using the appropriate sequencing primer listed in Table 3.1. All sequences were compared to a wildtype sequence generated from the reagent control cells to determine whether a knockout cell line was generated.

3.2.7 Analysis of the knockout status

DNA sequences generated in section 3.2.6 were analysed using Benchling and the Synthego ICE CRISPR Analysis Tool (https://ice.synthego.com/). Wildtype and potential knockout sequences were aligned using Benchling to identify whether an insertion or deletion (indel) had occurred. ICE was used to determine the percentage chance the cell line is a complete knockout through the 'Knockout Score'. Table 3.2 lists the indel that was generated for each cell line used in the experiments outlined in this thesis. A minimum of three individual cell lines with a knockout score greater than 90% was included. One cell line represents a cell clonally expanded from a single well from the single-cell sorting process (section 3.2.5). The use of three different clonal knockouts of a single gene enables more robust results to be generated as each knockout cell line functions as distinct biological replicates. Each unique cell line is represented by the 'cell line identifier' in Table 3.6. Generating an RPS4X knockout cell line was impossible, despite multiple *RPS4Y1* knockout cell lines being generated. The cell line identifier

Table 3.2: List of individual knockout cell lines used throughout this study. Indel indicates the type and size of the mutation, i.e. +1 = single base insertion and -1 = single base deletion. The knockout score represents the likelihood a knockout cell line was generated; the higher the score, the more likely.

Gene Knockout Target	Cell line identifier	Indel	Knockout Score
UTX	K6AC8	-16	99
	K6AF3	+1	99
	K6AG4	+1	100
UTY	K6CB3	-2	100
	K6CD8	-2	100
	K6CG11	-1	98
UTX/UTY Double-knockout (DKO)	ACB3	+1; +2	94
	ACB5	+1; -1	95
	ACB52	+1; -1	95
RPS4Y1	RYA2	-7	100
	FRCB3	-2	100
	RYF12	-2	100
ZFX	ZXA10	+1	99
	ZXB11	+1	99
	ZXD4	+1	99
ZFY	ZYC11	-4; -1	96
	ZYE7	-4	98
	ZY2B6	+1	99

3.3 Enzyme-Linked Immunosorbent Assay (ELISA)

Identification and quantification of CXCL8 and IL6 in the cell culture supernatant samples collected after 24-hour stimulations was performed using an enzyme-linked immunosorbent assay (ELISA), according to the manufacturer's instructions. All incubation steps were completed on an orbital shaker, and all washes were performed using (0.05% v/v) Tween20 - phosphate-buffered saline (T-PBS) (#0777, Amresco, OH, USA). Briefly, anti-CXCL8 capture antibodies (#554716, BD Pharmingen, NJ, USA) were diluted (1:500) in PBS, and anti-IL6 capture antibodies (#554543, BD Pharmingen, NJ, USA) were diluted (1:1000) in 0.1M disodium phosphate. Capture antibodies were incubated at 4°C overnight in 96-well immune plates (#NUN439454, ThermoFisher Scientific, MA, USA) at 100 µL/well. Plates were washed

before 1 hour of blocking with 1% (w/v) BSA-PBS at room temperature. During this incubation, serial dilutions were created for CXCL8 from 0 pg/mL to 1000 pg/mL (#554609, BD Pharmingen, NJ, USA) and IL6 from 0 pg/mL to 2000 pg/mL (#550071, BD Pharmingen, NJ, USA). After blocking, plates were washed once before the cell-free supernatant samples were loaded with the CXCL8 and IL6 standard dilutions. Samples were incubated at 4°C overnight. Following the incubation, plates were washed with T-PBS. CXCL8 (#554718, BD Pharmingen, NJ, USA) and IL6 (#554546, BD Pharmingen, NJ, USA) detection antibodies were incubated at room temperature for 1 hour. The plates were washed with T-PBS. Streptavidin horseradish-peroxidase (HRP) (#890803, R&D Systems, MN, USA) in T-PBS was incubated for 30 min in the dark at room temperature. At completion, the plates were washed. 3, 3', 5, 5' -tetramethylbenzidine (TMB) (#002023, Life Technologies, CA, USA) was added to each well for colour development and incubated for 2 – 3 min. The reaction was stopped by the addition of 1M phosphoric acid (H_3PO_4). Absorbance was measured at 450 nm and 570 nm for background using a Spectramax iD3 spectrophotometer.

3.4 Wound healing assay

Knockout cell lines were seeded in 12-well cell culture plates and incubated at 37°C/ 5% CO_2 until a cell monolayer was formed. Upon monolayer formation, a 200 µL pipette tip was used to smoothly "scratch" each well down the middle with even pressure. The media was aspirated, and the well was washed twice with HBSS. Growth media (1 mL) was gently added to the well, and a brightfield microscopy image was taken using a Nikon Eclipse Ti microscope. Cells were incubated at 37°C/ 5% CO_2 for 72 hours, with brightfield images taken every 24 hours.

Wound size was calculated by area using the Image-J plugin "Wound Healing Size". The analysis protocol as described by Suarez-Arnedo et al.[7] was followed. The wound size at time zero hours was considered 100%, with the size of the wound at subsequent timepoints calculated in comparison to zero hours.

3.5 Fibronectin-mediated cell adhesion assay

The protocol was adapted from what is described by Humphries 2000[8]. Fibronectin (#F0895, Sigma-Aldrich, Castle Hill, NSW) was diluted in sterile PBS, and 96-well plates were coated with fibronectin at 5 µg/cm^2 by incubation in sterile conditions for 1 hour at room temperature. The fibronectin solution was removed, and cells were seeded to both coated and uncoated wells at a density of one million cells per mL at 100 µL/well. A row on the 96-well plate was left empty to function as a no-cell control. The plate was agitated by semi-forceful tapping before incubation at 37°C/5% CO_2 for 1 hour. After the incubation, the plate was again tapped with even force to detach weakly adherent cells. Media was carefully aspirated, and the wells were gently washed once with sterile PBS. Adherent cells were fixed with 4%

paraformaldehyde (#29447, Merck-Millipore, MA, USA) by incubation at room temperature for 20 min. Afterwards, wells were washed twice with sterile PBS. Adherent cells were stained with 100 μL of 0.1% (w/v) crystal violet (#C6158, Sigma-Aldrich, Castle Hill, NSW), 200 mM MES (#M8250, Sigma-Aldrich, Castle Hill, NSW), pH 6.0 for 1.5 to 2 hours. Cells were inspected for staining using a brightfield microscope to determine when the dye had been completely absorbed. After staining, cells were washed with water three times before the dye was solubilised with 10% (v/v) acetic acid. The plate was agitated for 1 min to ensure complete solubilisation. Optical density (OD) was measured at 570 nm using the Spectramax iD3 plate reader. The background absorbance from the 'no-cell' lane was subtracted from all wells to determine cell attachment levels. Fibronectin-coated wells were compared to non-coated wells to determine the fold change in adherent cells.

3.6 Cigarette smoke extract (CSE) generation and treatment

In sterile conditions, 25 mL of unsupplemented DMEM was added to a T175 cell culture flask to generate CSE stimulation media. A 5 mL stripette was inserted through the flask filter with the tip submerged into the media with the flask upright, and an airtight seal was created using parafilm. A needle was inserted through the side of the flask and connected with tubing to a peristaltic pump. Tubing was also attached to the end of the 5 mL stripette, and a Marlborough Red cigarette was inserted into the end of the tubing and lit. The peristaltic pump was turned on to bubble cigarette smoke through the media. The speed of the pump was controlled to ensure the cigarette finished burning between 2 – 2.5 min. The process was repeated for a second cigarette. The media was allowed to settle for 5 min and supplemented with 0.1% (v/v) BSA. This was considered 100% CSE media.

For CSE exposure, all cell lines were seeded in a 96-well cell culture plate at $2x10^4$ cells/mL density and growth to 80% confluency. Cells were serum starved for 24 hours before exposure to CSE. A separate plate was created for non-treated cells to avoid incidental exposure to CSE components affecting baseline results. All cell lines were exposed to either 0%, 50%, 75% or 100% CSE media which were created from the original 100% CSE stock generated above. Cells were incubated at 37°C/ 5% CO_2 for 24 or 48 hours, with the MTT assay completed after the incubation period to assess cell viability. MTT methodology is described in section 3.7.

3.7 MTT assay

At the appropriate time point post-stimulation, 10 μL of 0.5% (w/v) MTT dye ((3-(4, 5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium) (#M2128, Sigma-Aldrich, Castle Hill, NSW) in PBS was added to each well of a 96-well plate under sterile conditions. All stimulations were completed in triplicate (to account for unavoidable experimental variability). Plates were incubated at 37°C/ 5% CO_2 for 4 hours. After the supernatant was removed and 100 μL of dimethyl sulfoxide (DMSO) was added to each well to dissolve the crystalline dye. After 1 min of vigorous shaking, plates were read on the Spectramax iD3 plate reader at 570 nm with a background reading at 630 nm. To analyse the data, the untreated well was considered to be 100% cell survival with all CSE-treated wells compared to the control well to determine the percentage of cell survival.

3.8 Extracellular matrix (ECM) ELISA

An ECM ELISA was performed to assess changes in ECM protein deposition between the knockout cell lines, as previously described[9]. Briefly, cell lines were seeded in 96-well plates and incubated in growth media at 37°C/ 5% CO_2 until confluency. Cells were serum starved for 18 hours before stimulation with TGF-β1 (10 ng/mL) for 72 hours. Cells were washed with PBS and lysed by incubation with 0.016 mM NH_2OH at 37°C for 15 min. The cell-free ECM was washed three times with sterile PBS. 100 μL of PBS was added to each well, with plates stored at -20°C until analysis. Before analysis, plates were defrosted at room temperature, with non-specific bound molecules removed via blocking with 1% BSA-PBS solution. Primary antibodies for collagen IVα1, tenascin-C, and fibronectin were diluted in 1X BSA-T-PBS and added at 100 μL/well for 2 hours. Plates were washed four times with T-PBS, and a rabbit anti-mouse monoclonal HRP-linked secondary antibody was applied for 1 hour. An IgG isotype control was also included for any non-specific IgG binding. Plates were washed, and 100 μL of chromogenic 3,3',5,5'-tetramethylbenzidine (TMB) was added for 15 – 30 min. Table 3.3 summarises the different antibody dilutions and incubation times used for each protein of interest. The reaction was stopped by the addition of 100 μL/well of 1M phosphoric acid. Absorbance was measured at 450 nm with background measured at 570 nm using a Spectramax iD3 plate reader.

Table 3.3: ECM ELISA primary and secondary antibodies with the dilution used and the time allowed for colourimetric development. IgG = immunoglobulin isotype G

Catalogue Number	Clone	Company	Antibody Target and Species	Antibody Dilution	Development time
MAB1935	868A11	Merck-Millipore	Fibronectin	1:4000	15min
T2551	BC-24	Merck-Millipore	Tenascin C	1:1000	30min
C1926	COL-94	Sigma-Aldrich	Collagen IV	1:1000	30min
557273		BD Pharmingen	IgG Isotype Control	1:1000	30min
7076S		Cell Signalling	Anti-mouse IgG-HRP-linked	1:2000	N/A

3.9 Cell Proliferation

Cell proliferation was investigated by measuring the increase in cell number after four days of growth. Cells were seeded in T75cm^2 flasks at a starting number of 400,000 cells. The flasks were incubated in growth media at 37°C/ 5% CO_2 for four days. After incubation, cells were harvested using trypsin-ETDA and counted using trypan blue exclusion on a haemocytometer. This number of cells was then compared to the starting seeding concentration. To calculate the doubling time of cell lines, the formula below was used:

$$Doubling\ time = \frac{96\ hours\ x \ln(2)}{\ln\left(\frac{Final\ cell\ count}{400,000}\right)}$$

3.10 RNA extraction and purification

Whole-cell mRNA extracts were collected and purified using the Bioline Isolate II Mini RNA Kit (#BIO52073, Meridian Bioscience, OH, USA), according to the manufacturer's instructions. Briefly, lysis buffer was supplemented with β-mercaptoethanol (β-ME) (#63689, Sigma-Aldrich, Castle Hill, NSW) to maximise cell lysis and mRNA stability. Cell lysis was completed mechanically, using a sterile scraper, after 350 µL of lysis buffer, supplemented with protease inhibitors, was added to the cell culture well. The lysate was transferred to a tube and vortexed briefly before being stored at -80°C until purification.

RNA lysates were defrosted slowly on ice. Samples were loaded to a supplied filter and centrifuged at 11,000g for 1 min to remove cell fragments and protein contaminants. The flow-through was mixed with 70% ethanol (made using RNase-free water), to adjust the nucleic acid binding conditions by mixing. Samples were loaded into the supplied spin column

containing a silica membrane filter that binds to nucleic acids. Samples were centrifuged at 11,000g for 30 sec. The flow-through was discarded, and membrane desalting was completed on the spin column using the supplied buffer and centrifugation. DNA was digested using DNase1 by incubation on the silica membrane for 15 min at room temperature. DNase1 was inactivated using the supplied wash solution before subsequent washing steps using the secondary wash buffer supplemented with >95% pure ethanol to remove contaminants and maximise RNA yield. RNA was eluted in 60 µL of RNase-free water. The NanoDrop 2000 UV-visible light spectrophotometer was used to analyse mRNA purity and quantity. Purified RNA samples were stored at -80°C until future use.

3.11 Bulk RNA-sequencing (RNA-Seq) processing and analysis

3.11.1 RNA-sequencing

The cell lines and stimulations sent for RNA-sequencing are listed below:

- Baseline: Wildtype, ZFX KO, ZFY KO, RPS4Y1 KO, UTX KO, UTY KO and UTX/UTY DKO
- 6 hours post-TNFα (10 ng/mL) stimulation: Wildtype, ZFX KO, ZFY KO & RPS4Y1 KO
- 48 hours post-TGF-β1 (10 ng/mL) stimulation: Wildtype, ZFX KO, ZFY KO, RPS4Y1 KO, UTX KO, UTY KO and UTX/UTY DKO

Bulk RNA-sequencing was completed by the Ramaciotti Centre for Genomics (University of New South Wales), with samples generated according to their instructions. Briefly, purified RNA samples were loaded into a clear 96-well hard-shell PCR plate. Total RNA was loaded into the plate at 50 ng/µL in a 50 µL volume. Only samples with a 260/280 ratio between 1.8 – 2.2 were included. Ramaciotti Institute determined the RNA integrity number (RIN), and a RIN between 7 – 10 was accepted as appropriate quality to continue with the analysis. Sequencing was completed using the polyA pull-down method, with single-ended reads for an average of 20 million reads per sample. Illumina stranded mRNA preparation ligation was completed before, and sequencing was conducted on a NovaSeq6000 S1 1x100bp flow cell.

3.11.2 Processing RNA-sequencing results

The quality of the RNA-seq samples was first checked using FASTQC. The paired-end reads were trimmed to remove adaptor sequences using Trimmomatic (v0.39). To estimate expression levels, single-end reads were mapped to human genome assembly GRCh38 (hg38) using Star aligner (v2.7.5a). The generated reads per gene counts were then used for differential gene expression, as described in section 3.12.

3.11.3 Quality assessment of RNA-sequencing

To assess the quality of the RNA-sequencing data, we measured the total counts determined for each sample and assessed whether it passed the threshold of five million total reads [10]. All but one sample passed this criterion (Figure 3.3A & 3.3B). Principle component analysis (PCA) was also completed to summarise which samples could be distinguished from others and to confirm whether samples were labelled accurately. The sample that did not have five million total reads was still grouped correctly in the PCA, therefore, this sample was still included. PCA analysis also revealed that a wildtype sample at baseline and after TNFα stimulation were mislabeled. As such, we corrected this labelling to ensure accurate differential gene expression analysis occurred. Figure 3.3C presents the PCA plot derived from the corrected samples, whilst Figure 3.3D presents the PCA plot for the 48-hour TGF-β1-stimulated samples.
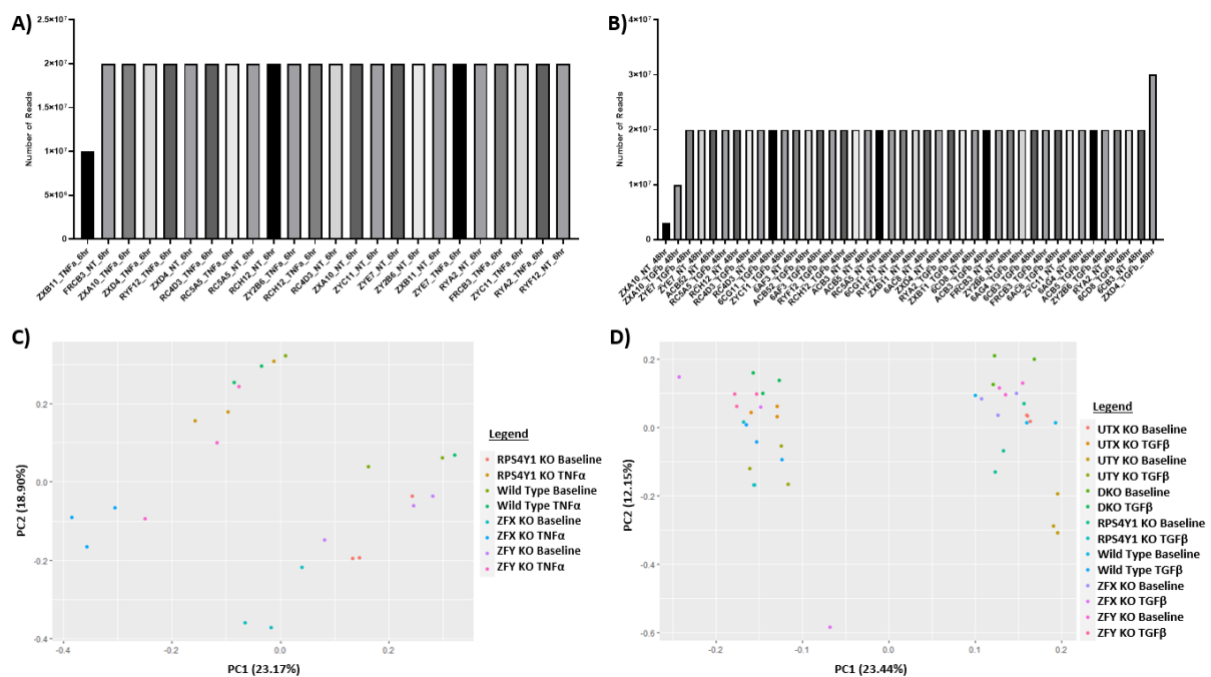
Figure 3.3: Quality assurance analysis of total read counts and principle component analyses (PCAs). Total number of reads for 6-hour (A) and 48-hour (B) models sent for RNA-sequencing; the x-axis presents the individual samples whilst the y-axis presents the total number of reads. PCA figures for 6-hour (C) and 48-hour (D) models plot principle component 1 (PC1) on the x-axis and PC2 on the y-axis, with the relative contribution included in brackets.

3.12 Differential gene expression analysis

Differential gene expression (DGE) was analysed using least-squares linear modelling in the R package DESeq2. Correction for multiple testing was applied by controlling the false discovery rate (FDR) using the Benjamini-Hochberg (BH) procedure [11]. Significant differentially expressed genes were determined with an FDR less than 0.05 and a fold-change (FC) greater than |1.0|. Standard DGE analysis involved comparing differences in gene expression between one cell genotype and another under the same conditions. For example, wildtype cells at baseline are compared to ZFX knockout cells at baseline. This would be written as the formula below:

$$(ZFX\ KO\ gene\ expression\ at\ baseline) - (Wildtype\ gene\ expression\ at\ baseline)$$

To evaluate the difference in responsiveness to TNFα and TGF-β1 between each cell line we completed an interaction differential gene expression analysis. Interaction DGE analysis involved comparing gene expression between baseline and stimulated cell lines of the same genotype to the change in gene expression of another cell genotype at baseline and stimulated. For example, the difference in gene expression of wildtype cells at baseline vs TNFα-stimulation, compared to the change in gene expression of ZFX knockout cells at baseline vs TNFα, this would be written as the formula below:

$$[(ZFX\ KO\ with\ TNF\alpha - ZFX\ KO\ at\ baseline)$$
$$- (Wildtype\ with\ TNF\alpha - Wildtype\ at\ baseline)]$$

3.13 Transcription factor enrichment analysis

We conducted a transcription factor analysis to analyse the signalling pathways that were specifically activated in each experimental condition for each cell line. This analysis generates an enrichment score based on the gene expression of known targets of transcription factors to represent whether that specific transcription factor pathway is more or less activated. To do this, ARACNe [12] (**A**lgorithm for the **R**econstruction of **G**ene **R**egulatory **Ne**tworks) (http://wiki.c2b2.columbia.edu/califanolab/index.php/Software/ARACNE) was used to identify transcription factors using a list of human transcription factors (n = 1639) downloaded from http://humantfs.ccbr.utoronto.ca/. One hundred bootstrap rounds were applied, and all inferred networks were merged to get a consensus network. Viper [13] was then applied to infer the protein activity of the identified transcription factor. Transcription factors with direct ARACNE-predicted T (regulon) statistically enriched in the gene expression signature were selected as the most representative transcription factors.

3.14 Gene set variation analysis (GSVA)

GSVA was used to investigate sample-wise gene set enrichment scores [14] to understand how the transcriptomic changes differ between knockout cell lines. A list of the top 50 unique differentially expressed genes that are either increased or decreased was generated to complete this. Using the R package '*GSVA*', the expression of the genes in this gene list was tracked for their change in replicates in another dataset. For example, the top 50 significantly upregulated genes in ZFX knockout cell lines at baseline were analysed for how their gene expression changed in ZFY knockout cell lines. This analysis produces an enrichment score, where numbers greater than zero indicate an overall upregulation, whilst an enrichment score less than zero indicates a downregulation. This highly versatile analysis was applied across multiple datasets described in the relevant sections.

3.15 Analysis of biological pathways altered in knockout cell lines

Pathway analysis was conducted to identify the biological pathways that involve the significantly differentially expressed genes in knockout cell lines compared to wildtype cells. The g: Profiler (https://biit.cs.ut.ee/gprofiler/gost) web-based tool was used to complete the analysis, using a subset of at most the top 50 significant genes (FDR < 0.05 and FC > |1.0|).

3.16 Independent single-cell and bulk RNA-sequencing study cohorts of asthma and COPD patients

Multiple independent studies were used throughout this thesis to support our investigation of *ZFX, ZFY, RPS4Y1, UTX* and *UTY*. The details of each study design, sample collection, protocols and clinical summary tables, where relevant, are included below.

3.16.1 Single-cell sequencing 'Lung Cell Atlas'

The publicly available single-cell RNA-sequencing database 'Lung Cell Atlas' (https://asthma.cellgeni.sanger.ac.uk/) was accessed to analyse the gene expression of *ZFX*, *ZFY*, *RPS4X*, *RPS4Y1*, *UTX* and *UTY* in different epithelial cell types from non-asthmatic and asthmatic patients. Patient samples were collected as part of a previous study by bronchoscopy biopsy. Complete details and methodology are published [15]. The cohort inclusion criteria for all patients were: aged between 40 – 65 years of age, < 10 pack years smoking history. The criteria for only the patients with asthma: age of onset < 12 years old, documented history of asthma, use of inhaled corticosteroids with(out) β_2-agonists due to respiratory symptoms, and positive provocation test (PC_{20} methacholine – induction of a 20% decrease in forced expiratory volume in the first second < 8 mg/mL with 2 min protocol). For non-asthmatic controls: absent history of asthma, no use of asthma-related medication, a negative provocation test (PC_{20} > 8 mg/mL, and adenosine 5'-monophosphate > 320 mg/mL with 2 min protocol), no pulmonary obstruction (FEV_1/FVC > 70%) and absence of lung

function impairment (FEV_1 > 80% predicted). Patients with asthma stopped inhaled corticosteroid use six weeks before testing and protocols were started. Patient demographics, as used in the current study, are detailed in Table 3.4.

Table 3.4: Demographics of patients included in the Lung Cell Atlas dataset.

Characteristic	Control	Asthma
n, (male / female)	7M / 4F	7M / 2F
Age	56	55
FEV_1 % Predicted	115.73 (+/- 13.86)	82.22 (+/- 19.56)
Pack years	1.27 (+/- 1.85)	0.00 (+/- 0.00)

FEV_1 = forced expiratory volume after one second. Data are presented as arithmetic mean (+/- SD).

3.16.2 Single-cell sequencing 'Human Lung Cell Atlas'

A single-cell dataset was obtained from the Human Lung Cell Atlas [16] and was prepared using the '*Seurat*' R package per the methodology outlined by Hao et al. [17]. The resulting, processed dataset was stratified by sex with the expression of *ZFX*, *ZFY*, *RPS4X*, *RPS4Y1*, *UTX,* and *UTY* visualised as a UMAP (uniform manifold approximation and projection for dimension reduction). The UMAP identified cell subpopulations such as epithelial cells, alveolar macrophages and immune cells. This enabled visualisation of which cell types the genes were expressed to provide insight into their expression patterns and function. A summary of the demographics for the patients included in the original data collection by Sikkema et al. [16] is summarised in Table 3.5.

Table 3.5: Demographics of patients included in the single-cell sequencing Human Lung Cell Atlas.

Characteristic	
n, (male / female)	64M / 43F
Age, range (yr)	10 - 76
BMI, (range)	20 - 49
Smoking Status, n	
Never smoker	56
Ex-smoker	17
Current smoker	16
Unannotated	18

BMI: body mass index

3.16.3 Indurian bulk RNA-seq bronchial biopsy dataset

Bronchial biopsies were collected from patients as part of a previous study [18]. The details of the study design and methodology are previously published [19]. All data is publicly available (EGA #337622). This study included persistent asthma subjects with clinical remission of asthma, complete remission of asthma and healthy control patients. We have only included patients with persistent asthma and healthy controls for the current study. The patient demographics are summarised in Table 3.6.

Table 3.6: Demographics table for the Indurian bronchial biopsy study patients.

Characteristic	Males	Females
Total, n (% Asthma)	88 (52.3%)	85 (58.8%)
Age, yr	46.46 (+/- 14.19)	41.29 (+/- 15.42) *
Pack years	10.42 (+/- 14.47)	5.27 (10.87) **
FEV_1, L	3.72 (+/- 0.91)	3.05 (+/- 0.68) ****
FEV_1 % Predicted	89.19 (+/- 16.13)	95.77 (+/- 14.59) **
Sputum cell counts Absolute (x10^3) % cell count		
Neutrophil	1.5 (+/- 2.33) 55.07% (+/- 19.35%)	1.68 (+/- 6.27) 53.42% (+/- 23.34%)
Eosinophil	0.04 (+/- 0.10) 1.82% (+/- 4.44%)	0.02 (+/- 0.09) 0.94% (+/- 1.80%)
Macrophage	1.05 (+/- 1.13) 42.45% (+/- 19.53%)	0.97 (+/- 1.43) 44.74% (+/- 22.92%)
Lymphocyte	0.02 (+/- 0.03) 0.66% (+/- 0.96%)	0.02 (+/- 0.03) 0.91% (+/- 1.37%)

*FEV_1 = forced expiratory volume after one second. Data are presented as arithmetic mean (+/- SD). Statistical significance is indicated by *p<0.05, **p<0.01, ****p<0.0001, determined by parametric t-test.*

3.16.4 OLiVIA clinical dataset patients and study design

Study subjects were selected from participants of the 'Effects **O**f Extra-fine Partic**L**e HFA-Beclomethasone **V**ersus Coarse Particle Treatment **I**n Smokers and Ex-smokers with **A**sthma' (OLiVIA study) (ClinicalTrials.gov #NCT01741285)[20]. This study was approved by the local medical ethics committee, with written and informed consent received before collection. Patients were included if they were either current- or ex-smokers (smoking cessation ≥ 6 months), aged 18 – 65 years old and had doctor-diagnosed asthma with a smoking history of ≥ 5 pack years. Exclusion criteria comprised treatment with oral steroids, $FEV_1 \leq 1.2$ L, an upper respiratory tract infection ≤ 4 weeks before inclusion, and an asthma exacerbation ≤ 6 weeks before inclusion. Nasal brushings were taken from the inferior turbinate of patients that met the inclusion criteria. RNA was isolated from the samples and quality tested before single-end sequencing. RNA sequencing data quality control was performed using R (version 3.4.3) to ensure concordance between reported sex and sex-associated genes. Data were $\log_2$ transformed and normalised. The patient demographics are summarised in Table 3.7.

Table 3.7: Clinical summary data for OLiVIA study patients.

	Never-smoker	Ex-smoker	Current-Smoker
n	18	30	54
Male, n (%)	10 (55.6)	13 (43.3)	26 (48.1)
Age, yr	50 (14.6)	50 (11.0)	41 (12.4) ** $
Pack years	0.00 (0.00)	21.10 (19.12)	20.77 (14.15)
FEV_1, L	3.68 (0.83)	2.91 (0.70) **	3.43 (0.84) $
FEV_1 % Predicted	107.4 (12.29)	90.00 (18.21) ***	95.06 (15.13) *

FEV_1 = Forced expiratory volume in one second (litres). Data are presented as the arithmetic mean +/- standard deviation and analysed using one-way ANOVA with Tukey correction for multiple testing. *$p<0.05$ (vs never smoker), **$p<0.01$ (vs never smoker), ***$p<0.001$ (vs never smoker), $ $p<0.05$ (current smoker vs ex-smoker).

3.16.5 SHERLOCk clinical dataset patients and study design

SHERLOCk (An integrative genomic approach to **S**olve t**H**e puzzle of sev**ER**e ear**L**y-**O**nset **C**OPD, ClinicalTrails.gov: NCT04263961 and NCT04023409) is a cross-sectional study without pharmacological intervention performed by the University of Groningen, the Netherlands. 23 non-COPD controls, 24 mild COPD (mCOPD) patients (GOLD stages 1 and 2) and 125 patients with severe COPD (sCOPD) (GOLD stages 3 and 4). Non-COPD was defined by an $FEV_1/FVC > 70\%$, mCOPD was defined as $30\% < FEV_1/FVC < 70\%$ and sCOPD was defined as $FEV_1/FVC < 30\%$. Participants did not smoke for at least two months before inclusion in the study and did not have an exacerbation or lung infection four weeks before the study. Subjects underwent bronchoscopy, during which bronchial and nasal brushes were obtained. All patients were fully characterized, i.e., lung function, CT scans, blood, and questionnaire data. The patient demographics are summarised in Table 3.8.

Table 3.8: SHERLOCk study patient clinical summary.

	Non-COPD	Mild-COPD (Stage 1 & 2)	Severe COPD (Stage 3 & 4)
N	23	24	125
Male, n (%)	12 (52)	18 (79)	36 (29)
Age, yrs	58 (+/- 5.4)	60 (+/- 4.6)	60 (+/- 7.0)
Ex-smoker, n (%)	23 (100)	23 (100)	124 (99)
Packyears	31.1 (+/- 20.56)	67.5 (+/- 61.69) ****	38.75 (+/- 18.16) *$$$$*
FEV_1/FVC (%)	74.0 (+/- 5.5)	53.9 (+/- 7.8) ****	28.6 (+/- 6.2) **** *$$$$*

FEV_1: forced expiratory volume in 1 second; FVC: forced vital capacity. Data are presented as the arithmetic mean +/- standard deviation and analysed using on-way ANOVA with Tukey correction for multiple testing. ****$p<0.0001$ (vs Non-COPD), *$$$$*$p<0.0001$ (severe COPD vs mild COPD smoker).

3.17 Mass-spectrometry proteomics analysis

3.17.1 Sample collection

Cell lines were seeded in 12-well cell culture plates as described in section 3.1 at 2.5×10^4 cells/mL. Cells were incubated in growth media until 80% confluent before serum starvation for 24 hours. Cells were then either stimulated with TNFα (10 ng/mL) for 24 hours or TGF-β1 (10 ng/mL) for 72 hours. At the end of the respective incubation periods, the supernatant was discarded, and cells were washed in ice-cold PBS twice before being lysed in sodium deoxycholate (SDC, #D6750, Sigma-Aldrich, Castle Hill, NSW) buffer supplemented with protein cocktail inhibitor III at 10 µL/mL (#539134, Sigma-Aldrich, Castle Hill, NSW) and phenylmethylsulphonyl fluoride (PMSF, #P7626, Sigma-Aldrich, Castle Hill, NSW) at 20 µL/mL. Physical lysis of cells was completed using a scraper to ensure maximal lysis and protein collection. Samples were stored at -80°C until the next step.

3.17.2 Alkylation and reduction of protein lysates

Samples from section 3.17.1 were defrosted on ice. Proteins in the sample were reduced and alkylated to denature secondary and tertiary structures to enable future digestion into peptides. To reduce proteins in the sample tris (2-carboxyethyl) phosphine (TCEP) at 5 mM was added and samples were vortexed briefly to mix. To alkylate samples, iodoacetamide (IAA) at 10 mM was added, and samples were vortexed briefly to mix. Samples were incubated at 70°C for 10 min and then cooled to room temperature. To deactivate reduction and alkylation components, dithiothreitol (DTT) was added at 10 mM, and samples were vortexed briefly and allowed to incubate for 5 min at room temperature in the dark. Total protein levels were quantified as described in section 3.17.3.

3.17.3 Bicinchoninic acid (BCA) assay

A BCA assay (#B9643, Sigma-Aldrich, Castle Hill, NSW) was completed to quantify the total protein amount in Mass spectrometry samples, following the manufacturer's instructions closely. Briefly, processed protein lysates generated in section 3.17.2 were vortexed to create a homogenous mixture. Samples were diluted by a factor of 40 to ensure absorbance values were within the range of the standard curve. The standard curve was generated using a filtered 10 mg/mL (w/v) BSA-PBS solution serially diluted across a concentration range of 0 – 1000 µg/mL. 25 µL of diluted sample and standards were loaded into a 96-well plate in duplicate. Appropriate BCA 'Reagent A' and 'Reagent B' volumes were mixed to form the BCA working reagent. 200 µL of this mixture was loaded into each well, and samples were incubated at 37°C for 30 min. Absorbance was measured at 562 nm, with the absorbance of the 0 µg/mL standard wells subtracted as 'background' absorbance from all other samples. Total protein amounts within samples were then interpolated from the standard curve.

3.17.4 Trypsin digest of protein samples

Using the protein concentration determined in section 3.17.3, 100 μg of protein was mixed with one μg of trypsin (#T1426, Sigma-Aldrich). Samples were incubated for 14 – 18 hours at 37°C to provide sufficient time for the digestion of proteins into peptides. Samples were processed immediately after incubation.

3.17.5 Solid phase extraction (SPE)

Digested peptide samples generated in section 3.17.4 were desalted and cleaned to remove factors introduced from enzymatic digestion. The protocol described here is adapted from Rappsilber et al. [21].

This method utilises styrenedivinylbenzene-reverse phase sulfonated (SDB-RPS, #2241, Empore, Sigma-Aldrich, Castle Hill, NSW) disks, which have been lodged within the end of a P200 pipette tip. These tips were placed through a hole in the top of an Eppendorf tube, so the tip would not touch the bottom of the tube. The constituents of each buffer used in this section are listed in Table 3.9. All centrifugation steps are completed at 5000g for 2 min. A 15 μg aliquot of the digested peptide was diluted in 10x the volume of SPE loading buffer, and the diluted protein samples were loaded into the pipette tip and centrifuged. During this step, the peptides bind to the SDB-RPS disc lodged in the pipette tip, and the flow-through is discarded. The SDB-RPS disk is washed with 100 μL of SPE loading buffer, and samples are centrifuged, with the flow-through discarded. Disks were washed again with 100 μL SPE wash buffer, and the flow-through was discarded. Before the elution of proteins from SDB-RPS disks, a mass-spectrometry loading vial was placed within the Eppendorf tube. The pipette tip was elevated approximately 1cm from the bottom (to avoid the flow through touching the end of the tip). To elute protein, 50 μL of SPE elution buffer was loaded into the pipette tip, and samples were centrifuged at 5000g for 5 min. Samples were then evaporated to dryness. Dry peptide samples were dissolved in 25 μL of loading solvent before being run on the mass spectrometer.

Table 3.9: SPE StageTip desalting and cleaning buffers. SPE = solid phase extraction; SPE elution buffer was made fresh every time

Buffer	Components % (v/v)
SPE Loading Buffer	90% Acetonitrile/ 1% Trifluoroacetic Acid
SPE Wash Buffer	10% Acetonitrile/ 0.1% Trifluoroacetic Acid
SPE Elution Buffer	80% Acetonitrile/ 5% Ammonium Hydroxide
Loading Solvent	2% Acetonitrile/ 0.1% Trifluoroacetic Acid

3.17.6 LC-MS/MS

Liquid chromatography – tandem mass spectrometry (LC-MS/MS) was completed to analyse the peptide abundance in each sample. Using an Acquity M-class nanoLC system (Waters, USA), 5 µL of the sample was loaded at 15 µL/min for 3 min onto a nanoEase Symmetry C18 trapping column (180 µm x 20 mm) before being washed onto a PicoFrit column (75 µmID x 350 mm; New Objective, Woburn, MA) packed with SP-120-1.7-ODS-BIO resin (1.7 µm, Osaka Soda Co, Japan) heated to 45°C. Peptides were eluted from the column and into the source of a Q Exactive Plus mass spectrometer (ThermoFisher) using the following program: 5-30% MS buffer B (98% Acetonitrile + 0.2% Formic Acid) over 90 min, 30 – 80% MS buffer B over 3 min, 80% MS buffer B for 2 min, 80-5% MS buffer B for 3 min. The eluting peptides were ionised at 2400V. A Data Dependent MS/MS (dd-MS^2) experiment was performed, with a survey scan of 350 – 1500 Da performed at 70,000 resolution for peptides of charge state 2+ or higher with an automatic gain control (AGC) target of $3x10^6$ and maximum injection time of 50 ms. The Top 12 peptides were selected and fragmented in the higher-energy C-trap dissociation (HCD) cell using an isolation window of 1.4 m/z, an AGC target of $1x10^5$ and a maximum injection time of 100 ms. Fragments were scanned in the Orbitrap analyser at 17,500 resolution, and the product ion fragment masses were measured over a mass range of 120 – 2000 Da. The mass of the precursor peptide was then excluded for 30 sec.

3.17.7 Mass spectrometry data analysis

The output LC-MS/MS data files from section 3.17.6 were searched using MaxQuant analysis software against the Uniprot human database (downloaded March 2021) with label-free quantification (LFQ). Search parameters used are as follows; Fixed modifications: None; variable modifications: oxidation, carbaminomethyl, deamination; Enzyme: semi-trypsin; the number of allowed missed cleavages: 3; peptide mass tolerance: 10 ppm; MS/MS mass tolerance: 0.05 Da. The results of the search were then inputted into the online proteomics visualization tool, LFQ-Analyst (https://analyst-suite.monash-proteomics.cloud.edu.au/apps/lfq-analyst//) [22]. From LFQ-Analyst the imputed raw data was

analysed using the R package *'limma'*. Differential protein analysis was conducted with significant differentially expressed proteins determined at an FDR < 0.05 and a fold-change > |1.0|. Multiple comparison testing was adjusted for using the Benjamini-Hochberg methodology [11].

3.18 Protein quantitative trait loci (pQTL) analysis

pQTL analysis was used to investigate how changes in gene expression correlate to changes in protein abundance in each cell line at baseline and after treatment with TNFα or TGF-β1. Gene expression at 6 hours was compared to protein abundance after 24 hours, whilst gene expression at 48 hours was compared to protein abundance after 72 hours. This analysis was completed in R, using the package '*MatrixEQTL*'. This package is designed to test for associations between two factors using linear regression with either additive or ANOVA gene expression effects. Normalised gene expression and protein values were generated, and all genes and proteins identified were used for the analysis. This technique identified the genes whose change in expression across cell lines correlated with the observed change in protein levels. Therefore, we could confirm that transcriptional regulation of these genes results in a functional change at a protein level. To account for multiple comparisons, the Benjamini-Hochberg method [11] of correcting for multiple comparisons was used with a false discovery rate (FDR) threshold of less than 0.05, to determine a statistically significant correlation. Once a correlation was identified, gene expression and protein levels were graphed using GraphPad Prism v9.0.

3.19 Western blotting

3.19.1 Sample collection

Western blots were used to visualise the abundance of ZFX, ZFY, RPS4Y1, UTX and UTY proteins as well as trimethylation status of histone 3 at lysine residue 27 (H3K27me3). Wildtype and knockout cell lines were seeded in 6-well cell culture plates at 200,000 cells/ml and incubated at 37°C/ 5% CO_2 in growth media until 100% confluency was reached, to maximise the amount of total protein collected. Supernatant was removed and cells were washed with ice cold PBS twice. Cells were lysed in protein lysis buffer including protease inhibitor cocktail and PMSF. Mechanical lysis of cells was ensured by the use of a sterile scraper and passing the lysate through a syringe and 25-gauge x 0.6" needle (#DN25GX0.6LV, Livingstone, NSW, AUS) five times. To analyse the abundance of H3K27me3 samples were processed and used immediately after lysis to minimise loss of methylation marks, via freeze thawing cycles [23]. All other samples were stored at -80°C, until analysis.

3.19.2 Preparation of polyacrylamide gels

Protein abundance was quantified in all samples prior to western blotting using BCA as described in section 3.17.3. Mini-PROTEAN ® 0.75 mm integrated spacer plates (#1653310, Bio-Rad, CA, USA) were used to cast western blot gels. 10% (v/v) acrylamide analysis gels were made by mixing reagents, as outlined in Table 3.10. 500 µl of 1-butanol (Cat# 537993, Sigma-Aldrich) was added to the centre of the casting plate to flatten the top edge of the gel and remove air bubbles. After the gel had set, butanol was removed, and the top of the gel was washed twice with milli-Q water. The 4% (v/v) acrylamide stacking gel mixture was made by mixing reagents as per Table 3.10. The mixture was added to the analysis gel with a ten-well plastic comb and allowed to set. Once completely set, gels were kept in an SDS-PAGE tank submerged in running buffer (Table 3.11) until samples were prepared.

Table 3.10: Components and volumes for making PAGE gel.

Reagent	10% (v/v) Acrylamide Analysis Gel	4% (v/v) Acrylamide Stacking Gel
10% (v/v) SDS/Tris Buffer (pH 8.8)	1.25 mL	-
10% (v/v) SDS/Tris Buffer (pH 6.8)	-	625 µL
40% Acrylamide	1.25 mL	250 µL
TEMED	5 µL	2.5 µL
10% (w/v) APS	50 µL	25 µL
Milli-Q Water	2.5 mL	1.63 mL

Table 3.11: List of buffers and reagents used to conduct western blotting.

Buffer	Components	pH
Cell lysis buffer (RIPA buffer)	150 mM NaCl (#S5886, Sigma-Aldrich) 1.0% (v/v) Triton X-100 (#X100, Sigma-Aldrich) 0.5% (w/v) SDC (D6750, Sigma-Aldrich) 0.1% (w/v) SDS (#L3771, Sigma-Aldrich) 50mM Tris-base, pH 8.0 (#BIO3094T, Astral Scientific)	N/A
5x loading buffer	0.312 M Tris-HCl (#, Sigma-Aldrich) 50% (v/v) Glycerol (#G6279, Sigma-Aldrich) 10% (w/v) SDS 0.5 M DTT (#D3483, Sigma-Aldrich) 0.05% (w/v) Bromophenol blue (#B0126, Sigma-Aldrich)	N/A
10x running buffer	0.25 M Tris-base 1.92 M Glycine (#BIOBG0235, Astral Scientific) 1% (w/v) SDS	N/A
Transfer buffer	0.478 M Tris-base 0.386 M Glycine 20% (v/v) Methanol (#MA004, ChemSupply, AUS)	N/A
Washing buffer (T-TBS)	200 mM Tris-base 1.5 M NaCl 0.05% (v/v) Tween-20 (#0777, Astral Scientific)	pH 7.2 – 7.4
Blocking buffer (Skim Milk-T-TBS)	5% (w/v) Skim milk powder (Woolworths, AUS) 200 mM Tris-base 1.5 M NaCl 0.05% (v/v) Tween-20	pH 7.2 – 7.4
Incubation buffer (BSA-T-TBS)	1% (w/v) BSA (#A4503, Sigma-Aldrich) 200 mM Tris-base 1.5 M NaCl 0.05% (v/v) Tween-20	pH 7.2 – 7.4

3.19.3 Electrophoresis

40 μg of protein was aliquoted from whole-cell protein lysates into a fresh Eppendorf tube and mixed with 5 μl of loading buffer (Table 3.10). The final volume of the sample/loading buffer mixture was made up to 25 μL with excess RIPA buffer. Mixed samples were heated at 95°C for 10 min to denature proteins and impart an overall negative charge. 20 μL of the denatured sample was loaded into a well in the polyacrylamide gel. 7 μL Precision Plus Protein ™ Dual Colour Standard protein ladder (Cat#161-0374, Bio-Rad, CA, USA) was loaded into one lane. Samples were electrophoresed in running buffer for 120 min at a constant voltage of 120 V using a Bio-Rad Powerpac 3000 system.

3.19.3 Protein transfer to PVDF membrane

After electrophoresis, the gel was removed from the cast, and the stacking gel layer was discarded. The remaining analysis gel was layered on top of both cellulose chromatography paper (#3030-917, GE Healthcare Life Sciences) and a 0.45 μm pore polyvinylidene difluoride (PVDF) membrane (# IPVH00010, Merk Millipore, MA, USA), pre-activated in 100% methanol for 2 min. To avoid contamination of protein, sterile and clean forceps were used to maneuver components. All layers were saturated in a transfer buffer (Table 3.11). Finally, another set of transfer buffer saturated chromatography paper was placed on top to complete the transfer sandwich. The transfer sandwich was placed within a plastic mould and placed inside a transfer tank. The tank was filled with transfer buffer with an ice block to maintain a cool temperature during the transfer process. Protein transfer was completed at 100 V for 1 hour.

3.19.4 Membrane blocking

After protein transfer, the PVDF membrane was incubated in blocking buffer (Table 3.11). Protein-containing membranes were blocked in 25 mL of blocking buffer for 45 min at room temperature on a rotating shaking platform.

3.19.5 Primary antibody incubation

After blocking, membranes were washed briefly three times with washing buffer (Table 3.11). Membranes were placed in a 50 mL Falcon tube, lining the inside so that the protein-containing side was facing towards the centre of the tube. 5 mL of incubation buffer containing the appropriate primary antibody at the required dilution (Table 3.12) was added to the 50 mL Falcon tube. Tubes were incubated on a roller tube tilt mixer platform at 4°C overnight.

3.19.6 Secondary antibody incubation

After primary antibody incubation, membranes were washed three times for 5 min in T-TBS buffer. The appropriate HRP-linked secondary antibody was diluted in 10 mL of 1% BSA-T-TBS buffer, as listed in Table 3.12. Membranes were incubated at room temperature for one hour.

3.19.7 Protein visualisation

After secondary antibody incubation, membranes were washed in T-TBS buffer four times for 5 min. Western Lightning Plus, Chemiluminescent Substrates (#NEL103E001EA, Perkin Elmer, MA, USA) were mixed in equal volumes in a tray, and the membrane was placed onto the mix with the protein side down. In the dark, the membrane was incubated in the substrate mixture for approximately 1 min. The membrane was placed protein side up on the imaging tray of a Bio-Rad ChemiDoc imager, using forceps. Proteins were visualised with UV exposure, with images captured in 10 sec intervals to avoid overexposure of the protein bands.

3.20 Statistical analysis

All results and figure legends indicate the statistical analysis, methods for the correction for multiple comparison testing, biological replicates, p-values and false discovery rates (FDR). Statistical significance was determined at a p-value or FDR less than 0.05. All error bars are presented as arithmetic mean +/- standard error of the mean (SEM). All comparisons were completed as two-tailed statistical tests. Dot plots were generated and analysed in GraphPad Prism v9.0. Volcano plots and heatmaps were generated using R software and the packages '*ggplot*' and '*heatmap3*'.

Table 3.12: List of antibodies used for western blotting protocol.

Catalogue Number	Clone	Company	Antibody Target (Species)	Antibody Dilution	Predicted Protein Size (kDa)
5419S	L28B6	Cell Signalling	ZFX (mouse)	1:250	135
PA5-68440	Polyclonal	Invitrogen	ZFY (rabbit)	1:200	100
17296-1-AP	Polyclonal	ProteinTec	RPS4Y1 (rabbit)	1:1,000	29
33510S	D3Q1I	Cell Signalling	UTX (mouse)	1:1,000	180
PA5-68440	Polyclonal	Invitrogen	UTY (rabbit)	1:200	160
61018	323	Biosearch	H3K27me3 (mouse)	1:1,000	17
MAB374	6C5	Merck-Millipore	GAPDH (mouse)	1:5,000	37
AP160P	N/A	Merck Millipore	Rabbit anti-mouse HRP linked	1:10,000	Secondary antibody
P0448	N/A	Dako	Goat anti-rabbit HRP linked	1:2,000	Secondary antibody

Chapter 4

4.1 Publication Declaration - 'Current Smoking Alters Gene Expression and DNA Methylation in the Nasal Epithelium of Patients with Asthma'

Karosham D. Reddy, Andy Lan, Ilse M. Boudewijn, Senani N.H. Rathnayake, Gerard H. Koppelman, Hananeh Aliee, Fabian Theis, Brian G. Oliver, Maarten van den Berge & Alen Faiz (2021), *American Journal of Respiratory Cell and Molecular Biology*.

Status: Published. The supplementary files are located in **Appendix A**.

Author Contributions: M.v.d.B. and A.F. conceived the idea. K.D.R., A.L., I.M.B., S.N.H.R., M.v.d.B., and A.F. contributed to, performed, and discussed data analysis and interpretation. H.A. and F.T. contributed to the data analysis. The National Heart, Lung, and Blood Institute LungMAP Consortium contributed to the data analysis. K.D.R., I.M.B., G.H.K., B.G.O., M.v.d.B., and A.F. discussed and contributed to the drafted manuscript composition.

Signatures:

Name:	Signature:	Date:
Karosham D. Reddy	Production Note: Signature removed prior to publication.	**08/12/2022**
Andy Lan	Production Note: Signature removed prior to publication.	**10/01/2023**
Ilse M. Boudewijn	Production Note: Signature removed prior to publication.	**23/01/2023**
Senani N.H. Rathnayake	Production Note: Signature removed prior to publication.	**23/01/2023**
Gerard H. Koppelman	Production Note: Signature removed prior to publication.	**23/01/2023**
Hananeh Aliee	Production Note: Signature removed prior to publication.	**01/03/2023**
Fabian Theis	Production Note: Signature removed prior to publication.	**11/01/2021**
Brian G. Oliver	Production Note: Signature removed prior to publication.	**08/12/2022**
Maarten van den Berge	Production Note: Signature removed prior to publication.	**23/01/2023**
Alen Faiz	Production Note: Signature removed prior to publication.	**23/01/2023**

ORIGINAL RESEARCH

Current Smoking Alters Gene Expression and DNA Methylation in the Nasal Epithelium of Patients with Asthma

Karosham D. Reddy[1,2], Andy Lan[3,4], Ilse M. Boudewijn[3,4], Senani N. H. Rathnayake[1], Gerard H. Koppelman[4,5], Hananeh Aliee[6], Fabian Theis[6], Brian G. Oliver[1,2], Maarten van den Berge[4*], and Alen Faiz[1,2,3,4*]

[1]School of Life Sciences, University of Technology Sydney, Sydney, Australia; [2]Respiratory Cellular and Molecular Biology, Woolcock Institute of Medical Research, University of Sydney, Sydney, Australia; [3]Department of Pulmonary Diseases, [4]Groningen Research Institute for Asthma and COPD, and [5]Beatrix Children's Hospital Department of Pediatric Pulmonology and Pediatric Allergology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; and [6]German Research Center for Health and Environment (GmbH), Helmholtz Centre, Munich, Germany

ORCID IDs: 0000-0001-9002-6930 (K.D.R.); 0000-0002-5631-7264 (S.N.H.R.).

Abstract

Current smoking contributes to worsened asthma prognosis and more severe symptoms and limits the beneficial effects of corticosteroids. As the nasal epithelium can reflect smoking-induced changes in the lower airways, it is a relevant source to investigate changes in gene expression and DNA methylation. This study explores gene expression and DNA methylation changes in current and ex-smokers with asthma. Matched gene expression and epigenome-wide DNA methylation samples collected from nasal brushings of 55 patients enrolled in a clinical trial investigation of current and ex-smoker patients with asthma were analyzed. Differential gene expression and DNA methylation analyses were conducted comparing current smokers with ex-smokers. Expression quantitative trait methylation (eQTM) analysis was completed to explore smoking-relevant genes by CpG sites that differ between current and ex-smokers. To investigate the relevance of the smoking-associated DNA methylation changes for the lower airways, significant CpG sites were explored in bronchial biopsies from patients who had stopped smoking. A total of 809 genes and 18,814 CpG sites were differentially associated with current smoking in the nose. The *cis*-eQTM analysis uncovered 171 CpG sites with a methylation status associated with smoking-related gene expression, including *AHRR, ALDH3A1, CYP1A1, and CYP1B1*. The methylation status of CpG sites altered by current smoking reversed with 1 year of smoking cessation. We confirm that current smoking alters epigenetic patterns and affects gene expression in the nasal epithelium of patients with asthma, which is partially reversible in bronchial biopsies after smoking cessation. We demonstrate the ability to discern molecular changes in the nasal epithelium, presenting this as a tool in future investigations into disease-relevant effects of tobacco smoke.

Keywords: asthma; nasal epithelium; gene expression; DNA methylation; smoking

The long-term adverse health effects of cigarette smoking are well established. Short-term tobacco smoke exposure negatively affects lung function, whereas long-term exposure is linked with chronic respiratory disease (1). Cigarette smoking increases the risk of asthma development and severity (2). Asthma causes 450,000 deaths worldwide annually, and 350 million people have it (3). The smoking rate among patients with asthma is similar to that of the general population, and they are less likely to quit (2). Smoking patients

with asthma are a complex and vulnerable population because they are more susceptible to the development of respiratory diseases such as chronic obstructive pulmonary disease (COPD), in addition to a poor asthma prognosis (4). Smoking patients with asthma experience more life-threatening exacerbations, making disease management difficult. The beneficial effect of inhaled corticosteroids on lung function outcomes is negated by smoking (5). As such, an altered physiological response occurs in the airways because of smoke exposure, causing worsened asthma disease outcomes (2).

The respiratory airway epithelium is the body's first line of defense against inhaled toxic agents and exhibits smoking-associated changes (6). Although bronchial samples are considered the most representative (7), the nasal epithelium can reflect smoking-induced changes in the lower airways. Functionally similar genes in paired nasal and bronchial epithelium samples can differentiate between healthy never-smokers and current smokers (7–11). Although nasal brushings do not perfectly mirror the lower airways, the response to tobacco smoking is shown to be reflective of transcriptional changes, and they are an attractive alternative to invasive bronchial biopsies.

Epigenetics refers to chemical changes to the genome that do not alter the nucleotide sequence (12). DNA methylation is the most widely studied epigenetic mark and is known to influence gene expression. Modifications in DNA methylation can be caused by environmental stimuli, such as cigarette smoke (13, 14). These changes can persist and contribute to increased disease susceptibility (15). As smoking increases asthma severity and reduces treatment effectiveness, greater understanding of gene expression and DNA methylation changes due to smoke exposure is needed. Methylation commonly occurs at positions in the DNA where a cytosine nucleotide is followed by a guanine nucleotide, these locations are referred to as CpG sites. Differentially methylated CpG sites exist between current, ex-, and never-smokers in samples taken from the lung and blood tissues. These sites are associated with the genes *ALDH3A1, CYP1A1, CYP1B1*, and *AHRR* (16, 17). Most studies investigate DNA methylation with blood samples; however, examination of structural cells is vital because they are exposed to smoke. Some studies investigate changes in gene expression and DNA methylation using patient-matched lung samples (11, 18, 19) but are limited by size. These analyses enable a direct investigation between gene expression and

DNA methylation, increasing our understanding of molecular processes that influence the tobacco smoke response.

The current study investigates nasal differential gene expression and DNA methylation profiles between current and ex-smokers with asthma. We examined differential gene expression due to current smoking in two *in vitro* datasets and in two *in vivo* smoking cessation datasets. We tested these gene expression patterns for an association with relevant clinical measures in asthma. Analysis of DNA methylation and gene expression using expression quantitative trait methylation (eQTM) analysis offered insight into the relationship between DNA methylation status and gene expression relative to smoking status. To address the dynamics of smoking-associated changes in the airways, significant nasal CpG sites were explored in bronchial biopsies from patients who had stopped smoking. This study contributes to the current understanding of nasal gene expression as an alternative to invasive bronchial biopsies. It elucidates mechanisms and identifies appropriate biomarkers affecting processes in smoking patients with asthma.

Methods

Participants and Sample Collection
Study subjects were selected from participants of the OLiViA (Effects of Extra-Fine Particle HFA-Beclomethasone Versus Coarse Particle Treatment in Smokers and Ex-Smokers with Asthma) study (ClinicalTrials.gov #NCT01741285) (20). This study was approved by the local medical ethics committee, with written and informed consent received before collection. Patients were included if they were either current or ex-smokers (smoking cessation $\geq$6 mo), aged 18–65 years old, and had doctor-diagnosed asthma with a smoking history of 5 pack-years or more. Exclusion criteria comprised treatment with oral steroids, forced expiratory volume in 1 second (FEV_1) of $\leq$1.2 L, and an upper respiratory tract infection $\leq$4 weeks before inclusion, had used inhaled corticosteroids in the previous 4–6 weeks, and an asthma exacerbation $\leq$6 weeks before inclusion. A schematic representation of the study design and analysis is included in the supplementary file (Figure E1 in the data supplement).

Sample Collection, Processing, and Quality Assessment
Detailed methods are presented in the data supplement. Nasal brushings were taken from

the inferior turbinate of patients. Two separate brushings were taken for RNA and DNA collections. RNA was isolated from samples using an RNeasy kit (#74106, Qiagen) and quality tested before single-end sequencing. Quality control of gene expression was performed using R, with concordance between reported sex and sex-associated genes (*XIST* and Y-chromosome genes) confirmed. Genes with mean expression of $\leq$1 fragments/million across all samples were removed. The influence of technical variation was assessed via principal component analysis using the R package EDASeq. The variation in library sizes was explored for gene expression among samples by making a relative log expression plots. The library size cut-off ensured that sequences would be larger than one-quarter of the median of all library sizes. Data were $\log_2$ transformed and normalized.

DNA extraction was completed on a second nasal brushing using a DNA extraction kit (#51306, Qiagen), according to the manufacturer's instructions. DNA methylation processing was completed using the Infinium MethylationEPIC BeadChip array. DNA methylation data quality and normalization were also assessed using the R packages *MethylAid* and *watermelon* with five diagnostic filter variables with the following thresholds: MU = 10, OP = 12, BS = 11.50, HC = 12.75, and DP = 0.95 (the definitions of these variables are provided in the data supplement). Discordant samples were discarded as "swapped" or "contaminated" samples. Probes with high intensities indistinguishable from the background, cross-reactive probes, and type I probes with high-intensity signals were discarded. Probes were excluded if intensities *1*) were indistinguishable from background levels, *2*) were cross-reactive, or *3*) were type I probes displaying high-intensity signals. The final number of CpG sites tested after filtration was 797,128 sites. Normalization was completed using the *dasen* method in the *watermelon* package, in which type 1 and type 2 background intensities are equalized separately using quantile normalization before the generation of β values. Detailed sample processing methodology is included in the data supplement.

Smoker versus Ex-Smoker: Differential Gene Expression Analysis
Differential expression was assessed using least-squares linear modeling in the R package limma. We corrected for age and sex by including each factor as a covariate in the linear model, whereas smoking status was the tested

variable. Correction for multiple testing was applied through controlling the false discovery rate (FDR) using the Benjamin-Hochberg (BH) procedure. Significant differentially expressed genes were determined with an FDR of less than 0.05 and a fold-change (FC) greater than |1.0|.

Cellular Deconvolution Analysis

Cellular deconvolution was completed using the Cibersort method on single-cell RNA sequencing signatures for basal, ciliated, and mucus-secretory cells (club and goblet cells) to investigate the composition of nasal brushings. Cell composition was then compared between current and ex-smokers. The complete methodology is included in the data supplement.

Gene Expression Profiling and Mapping with *In Vitro* Datasets

Independent publicly available *in vitro* datasets on the Gene Expression Omnibus were used to conduct gene set expression analysis (GSEA) and gene set variation analysis (GSVA). GSEA used datasets analyzing airway epithelial cells from patients exposed to cigarette smoke for 30 minutes four times (GSE30660) before gene expression analysis and bronchial epithelium air–liquid interface (ALI) cultures exposed cigarette smoke for 24 hours before gene expression analysis (GSE82137) (21). We conducted GSVA on a dataset that analyzed nasal epithelial gene expression from smokers who had quit smoking, over a 6-month period (GSE83364) (22). We also conducted GSVA on datasets that analyzed changes gene expression in A549 cells in which NRF2 (GSE113519) and AHR (GSE109576) (23) are inhibited. The methodology and patient demographics (where appropriate) are located in the data supplement.

Identification of Biological Pathways Altered by Smoke Exposure in the Nose

Pathway analysis was completed to identify the biological pathways that involve significant differentially expressed genes in nasal epithelium. The g:Profiler web-based tool was used to conduct the analysis, using a subset of the top 25 significant (FDR < 0.05 and FC > |1.0|) genes.

Association between Gene Expression Variation and Clinical Parameters

The top 25 significantly increased genes from the differential expression analysis were analyzed by GSVA to generate gene variation enrichment scores for each patient. This "smoking score" was analyzed for correlation with multiple clinical parameters, including pack-years, age, lung function (FEV_1/forced vital capacity), and peripheral immune cell counts. Data were analyzed using a nonparametric Spearman's correlation test, with statistical significance determined at $P < 0.05$.

Smoker versus Ex-Smoker: Differential DNA Methylation Analysis

Differential DNA methylation analysis used robust linear modeling in the R package limma. To remove the dependence of the mean (heteroscedasticity), β values were logit-transformed to *M* values. Smoking status, age, and sex were covariates in the linear model. Major principal components, explaining 95% of the variation in technical control probes, were included as covariates to account for variation, as recommended by Lehne and colleagues (24). The Δβ values were calculated by subtracting the average β value per CpG site for the current smoker group from the average β value for the ex-smoker group. Correction for multiple testing was controlled at FDR < 0.05 using the BH procedure. The BH correction method was preferred to the Bonferroni method to ensure consistency across analyses; because BH is the standard method for differential gene expression, it was used for DNA methylation analysis. As such, DNA methylation analysis was less restrictive to enable an appropriate comparison with gene expression. Bonferroni significance was still indicated in the figures for reference.

Expression Quantitative Trait Methylation Analysis

cis-eQTM analysis was performed using the R package MatrixEQTL (version 2.2). Age and sex were entered as covariates in the model. For the analysis, a 1-Mb region on either side of the gene was selected. The analysis was completed for each significant gene from the differential gene expression analysis (FDR < 0.05 and FC > |1.0|) individually against significant CpG sites from the differential DNA methylation analysis (FDR < 0.05). Correction for multiple testing was completed using the BH procedure.

Mapping of DNA Methylation Sites with Transcription Factor Binding Sites

The relevance of significant differential DNA methylation sites was profiled by comparison with independent online chromatin IP sequencing (ChIP-Seq) datasets. GSE90550 investigates the binding locations of smoking-related transcription factors AHR and AHRR in MCF-7 cells treated with DMSO or 2,3,7,8-Tetrachlorodibenzodioxin (TCDD) for 24 hours (25), whereas GSE75812 investigates transcription factor NRF2 binding locations in BEAS-2B cells treated with sulforaphane for 5 hours (26). The location of transcription binding sites was aligned with gene transcript locations and significant differentially methylated CpG sites identified in the eQTM analysis.

Longitudinal Analysis of Changes to DNA Methylation Sites with Smoking Cessation

The longitudinal effects of smoking cessation were evaluated using an independent 1-year smoking cessation model (27). The independent study completed differential DNA methylation analysis on paired bronchial biopsies taken from smokers with COPD before and after they had stopped smoking for 1 year. From a total of 63 participants, only 19 patients had successfully ceased smoking by the end of the study. Significant *cis*-eQTM CpG sites identified in the nasal epithelium as part of the current study were extracted from the differential DNA methylation results of the smoking cessation analysis. We investigated how the DNA methylation sites affected by current smoking the nasal epithelium are altered in bronchial samples upon 1 year of smoking cessation.

Statistical Analysis

Differential gene expression, differential DNA methylation, and eQTM analyses were all conducted using R software (version 4.0.1). All other analyses were performed using the GraphPad Prism software (version 8), with statistical significance determined at $P < 0.05$.

Results

Patient Demographics and Sample Processing

Ten samples were removed after mRNA quality checking processes, leaving a final number of 55 high-quality samples for analysis. Gene expression data were collected from these brushings. After processing, 31 samples had sufficient quality DNA to undergo DNA methylation analysis. Finally, 29 samples remained in which both gene expression and DNA methylation data were

available for eQTM analysis. Patient information and demographics are summarized in Table 1.

Differential Gene Expression in Nasal Epithelium between Current and Ex-Smokers

Differential gene expression analysis identified 809 genes being upregulated or downregulated in the nasal epithelium by current smoking. In total, 153 genes showed increased expression and 656 showed decreased expression (FDR < 0.05 and log-fold change > |1.0|; Figure 1A). Using hierarchy clustering, the majority of smokers clustered separately from the ex-smokers. Table E1 lists the top significantly differentially expressed genes. Among the significantly upregulated genes in active smoker nasal epithelium were *CYP1A1*, *CYP1B1*, *ALDH3A1*, *NQO1*, and *AHRR*. The most significant downregulated genes included *SAA1*, *SAA2*, *CYP4X1*, *COL9A2*, and *CHST4* (Figure 1B). Of particular note are genes *CYP1A1* and *CYP1B1*, which were greatly upregulated in current smokers. with a log-fold change of 7.44 (FDR = 5.80×10^{-19}) and 4.31 (FDR = 8.19×10^{-14}), respectively (Figure 1B).

Cellular Composition of Nasal Brushings Are Similar between Current and Ex-Smoker Patients

To assess the uniformity of nasal samples and whether epithelium from current and

ex-smokers differs, we conducted cellular deconvolution using the support vector regression (SVR) and non-negative least squares (NNLS) analytical methods. The following four cell types were most abundant across the samples: goblet cells, basal epithelium, dendritic cells, and ciliated epithelium. Both techniques reported minimal variability in cell composition between samples taken from current and ex-smokers (Figure E2). Therefore, we observed no difference in sample cell composition from nasal brushings between current and ex-smokers.

Smoking Affected Genes in the Nose Are Similarly Altered in Independent *In Vitro* Analysis of Airway Cells

To investigate whether differences in gene expression were caused by changes in the nasal epithelium or to inflammatory cells, we analyzed *in vitro* datasets of airway epithelial cells exposed to gaseous tobacco smoke. Using GSEA, we investigated whether the 809 genes altered by current smoking in nasal epithelium are similarly affected by cigarette smoke in independent ALI cell culture datasets. The genes *ALDH3A1*, *AHRR*, *CYP1A1*, *CYP1B1*, *CYP1B1-AS1*, and *SLC7A11* were core enriched among the genes that had increased expression by smoke exposure in the human epithelial ALI dataset (GSE30660; Figure 2A) and the bronchial epithelial ALI dataset (GSE82137; Figure 2B). The genes *CYP4X1*, *COL9A2*,

SAA1, and SAA2 were core enriched among genes downregulated by smoke exposure in the ALI datasets (Figures 2A and 2B). In Figures 2A and 2B, enriched genes are represented by black vertical lines indicating the amount these differentially expressed genes from the dataset are overrepresented in nasal epithelium gene expression. Core enriched genes, represented by red vertical lines, are genes that are highly represented in a gene set, indicating that they are highly likely to be associated with the stimulus. This shows that gene expression changes caused by smoke exposure in the nasal epithelium align with independent findings, indicating that gene expression occurring in the upper airways in response to cigarette smoke is reflected in airway epithelial cells in an *in vitro* setting.

AHR and NRF2 Biological Signaling Pathways Are Involved in the Nasal Response to Smoking

Next, to determine which pathways are enriched in our analysis, we conducted a pathway analysis using g:Profiler. Here, we identified five pathways, including general pathways such as response to external stimulus and metabolism of xenobiotics together with well-known smoking-related pathways such as the AHR pathway and the NRF2 pathway (Table 2).

To investigate these pathways, we analyzed two publicly available datasets that

Table 1. Clinical Characteristics of the Study Population

Characteristic	Differential Gene Expression Analysis		Differential Methylation Analysis		eQTM Analysis	
	Current Smoker	Ex-Smoker	Current Smoker	Ex-Smoker	Current Smoker	Ex-Smoker
N	26	29	14	17	13	16
Female, %	53.85	55.17	53.85	55.17	46.15	56.25
Age, yr	40 ± 11	50 ± 10*	38 ± 11	50 ± 11[†]	38 ± 12	49 ± 11*
Pack-years	22 ± 13	21 ± 19	18 ± 10	17 ± 10	19 ± 9	17 ± 10
Daily cigarettes	17.68 ± 10.54	—	17.68 ± 10.54	—	18.15 ± 7.18	—
FEV$_1$% predicted	86.08 ± 13.75	89.34 ± 18.17	82.14 ± 12.45	85.06 ± 17.97	82.38 ± 10.03	87.69 ± 16.81
FEV$_1$/FVC, %	83.87 ± 11.73	82.64 ± 13.82	82.08 ± 10.21	80.86 ± 12.29	82.88 ± 11.16	81.88 ± 10.80
Peripheral blood cell absolute count, $\times 10^3$ cell/μl						
Leukocytes	9.14 ± 2.15	6.64 ± 1.50[‡]	9.14 ± 2.15	6.64 ± 1.50[†]	9.10 ± 2.19	6.39 ± 1.72[†]
Neutrophils	5.21 ± 1.23	3.74 ± 1.57[†]	5.21 ± 1.23	3.74 ± 1.57[†]	5.36 ± 1.65	3.65 ± 1.47[†]
Lymphocytes	2.80 ± 0.81	2.04 ± 0.46[‡]	2.80 ± 0.81	2.04 ± 0.46[†]	2.68 ± 0.81	1.91 ± 0.33[†]
Monocytes	0.70 ± 0.17	0.53 ± 0.16[†]	0.70 ± 0.17	0.53 ± 0.16[†]	0.68 ± 0.15	0.51 ± 0.16[†]
Eosinophils	0.36 ± 0.20	0.27 ± 0.18	0.36 ± 0.20	0.27 ± 0.18	0.32 ± 0.18	0.26 ± 0.19

Definition of abbreviations: eQTM = expression quantitative trait methylation; FEV$_1$ = forced expiratory volume in 1 second; FVC = forced vital capacity.
Data are presented as arithmetic mean ± SD and analyzed using a Mann-Whitney test.
*$P < 0.05$ (current smoker vs. ex-smoker).
[†]$P < 0.001$ (current smoker vs. ex-smoker).
[‡]$P < 0.0001$ (current smoker vs. ex-smoker).

72
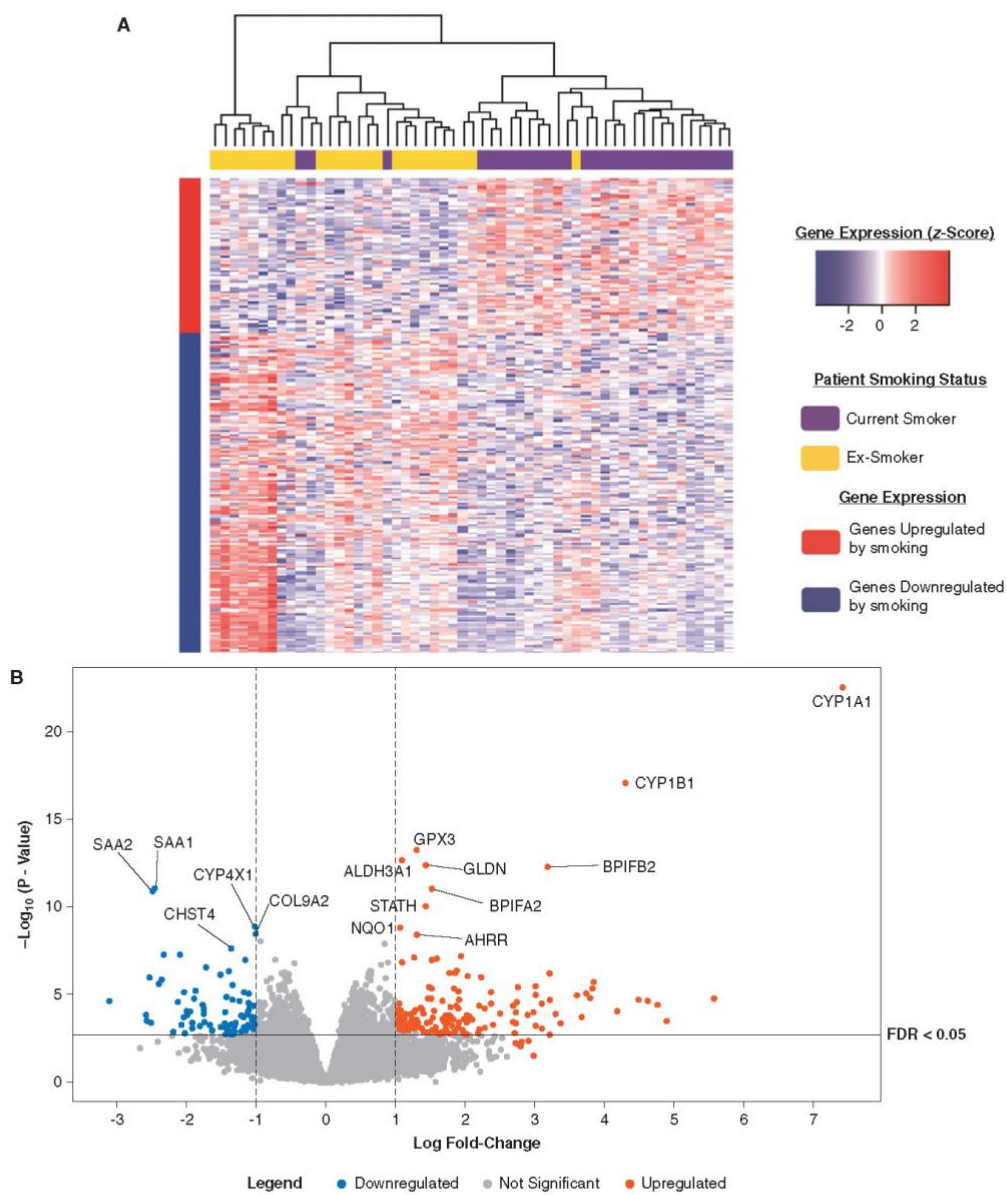
Figure 1. Heatmap and volcano plot visualization of differential gene expression between current smokers and ex-smokers. (*A*) An unsupervised heatmap of the 809 significant genes differentially expressed between smokers (purple) and ex-smokers (gold). Increased genes are colored red, and decreased genes are colored blue. (*B*) Volcano plot of $-\log_{10}$ (*P* value) against log-fold change in gene expression in smokers compared with ex-smokers. Significantly increased and decreased expressed genes in smokers are indicated by red and blue, respectively. Dotted vertical lines indicate fold-change of |1.0|. Statistical significance was determined at FDR-adjusted *P* value of less than 0.05, determined using the Benjamini-Hochberg method. $n = 26$ (current smokers) and 29 (ex-smokers). AHRR = aryl-hydrocarbon receptor repressor; ALDH3A1 = aldehyde dehydrogenase 1 family, member A3; CYP1A1 = cytochrome P450, family 1, subfamily A, polypeptide 1; CYP1B1 = cytochrome P450, family 1, subfamily B, polypeptide 1; FDR = false discovery rate.
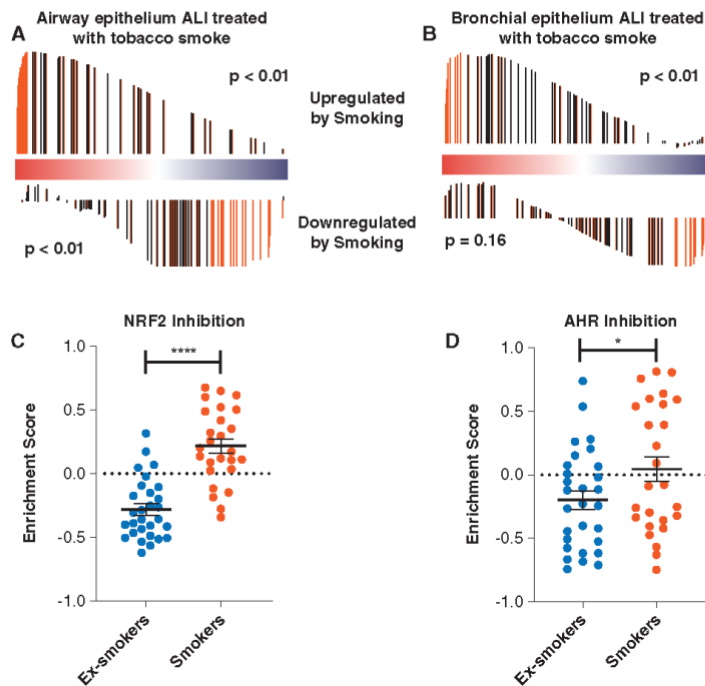
Figure 2. Analysis of smoking affected genes expressed in the nasal epithelium within independent *in vitro* datasets. (*A*) Gene set expression analysis (GSEA) enrichment of genes upregulated and downregulated in air–liquid interface (ALI) cultures of primary airway epithelial cells exposed cigarette smoke compared with nasal epithelium genes from patients with asthma. (*B*) GSEA enrichment of genes increased or decreased in bronchial epithelial ALI cultures upon cigarette smoke exposure compared with nasal epithelium genes from patients with asthma. Colored horizontal bars represent genes identified in each independent dataset ranked by *t* value; red indicates a positive association with gene expression, whereas blue indicates a negative association. The black vertical lines represent differentially expressed genes between smokers and ex-smokers in the nasal epithelium. The height of the black lines indicates the running enrichment scores of each gene, signifying the amount that the genes are overrepresented in the nasal epithelium. Red bars indicate core enriched genes; these are genes highly represented in the gene set. (*C* and *D*) Gene set variation analysis analyzing genes that indicate decreased expression in A549 cells upon NRF2 (*C*) or AHR inhibition (*D*) in A549 cells have increased expression in current smoker (red) nasal epithelial compared with ex-smokers (blue). Dots represent individual patients from nasal differential gene expression analysis. Data are presented as mean ± SEM and analyzed using Student's parametric *t* test. Statistical significance between ex-smoker and current smoker groups is indicated by *$P < 0.05$ and ****$P < 0.0001$. $n = 26$ (current smokers) and 29 (ex-smokers) for all analyses. AHR = aryl-hydrocarbon receptor.

specifically inhibited NRF2 (GSE113519) or AHRR (GSE109576) in airway epithelial cells. To understand the function of differentially expressed genes from these datasets in nasal epithelium from smokers with asthma, GSVA analysis was conducted. Because of limited genes indicating increased expression (FDR < 0.05) in these datasets, we focused on gene expression decreased by NRF2 and AHR inhibition. Genes that decreased in expression upon inhibition of either protein demonstrate an increased amount of expression in current smoker nasal epithelium (Figures 2C and 2D). These results indicate that both NRF2 and AHR have a significant role in the regulation of genes associated with current smoking.

Gene Expression Levels Reverse with Smoking Cessation

To explore the effect of smoking cessation on genes that change with smoke exposure, we performed a GSVA using an independent *in vivo* nasal smoking cessation dataset (22). A subset of the 25 top significantly increased and a second subset of the 25 top significantly decreased genes in the nasal epithelium were analyzed for changes associated with smoking cessation. Gene expression was analyzed 4, 8, 16, and 24 weeks after cessation. After 4 weeks, genes upregulated by current smoking in the nasal epithelium are downregulated, which is sustained over 24 weeks (Figure 3A). Comparatively, genes downregulated by current smoking in the nasal epithelium are upregulated 8 weeks after smoking cessation, which is sustained after 16 weeks (Figure 3B); at 24 weeks, this was no longer significant. It is unclear what might be the physiological effect of this change.

GSVA Smoking Score Correlates with Clinical Parameters

After this, we investigated whether transcriptional signatures were associated with clinical parameters (measures of lung function and peripheral inflammation). The subset of the top 25 significantly increased genes (FDR < 0.05 and FC > 1) in the nasal epithelium was used to generate a GSVA enrichment "smoking score." Figure 4 shows a positive correlation with our gene expression signature and pack-years ($P = 0.0077$). A significant correlation was also observed for peripheral blood cells such as leukocytes ($P < 0.0001$), neutrophils ($P = 0.0012$), monocytes ($P = 0.0015$), and lymphocytes ($P = 0.0010$); these have been included in the data supplement (Figure E3). However, when the correlation analysis is completed within each group, no significant correlation is observed for either current or ex-smokers ($P > 0.05$). No correlation was observed with age, FEV$_1$/forced vital capacity, number of cigarettes per day, or absolute eosinophil counts. No correlation is reported within the current smoker group for the clinical parameters against the GSVA smoking enrichment score, despite a correlation observed for pack-years. This might be explained by a lack of a broad range of smoking levels within the current smoker group. Therefore, the inclusion of a greater range of smoking levels will offer more nuanced insight between the ability to measure the effects of smoking on clinical parameters using this methodology.

Table 2. Summary of g: Profiler Pathway Analysis Investigating the Top 20 Significantly Upregulated Genes by Current Smoking

Pathway Name	Term ID	FDR
Cellular response to xenobiotic stimulus	GO:0071466	7.98×10^{04}
Metabolism of xenobiotics by cytochrome P450	KEGG:00980	8.53×10^{03}
Aryl Hydrocarbon receptor pathway	WP:WP2873	5.20×10^{07}
Nuclear receptors meta-pathway	WP:WP2882	1.43×10^{05}
NRF2 pathway	WP:WP2884	5.21×10^{03}

Definition of abbreviations: FDR = false discovery rate; GO = Gene Ontology; ID = identification; KEGG = Kyoto Encyclopedia of Genes and Genomes; WP = WikiPathways.
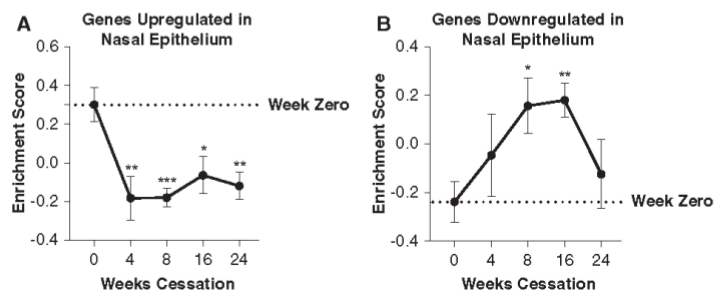
Figure 3. Gene set variation analysis demonstrating gene expression reverses upon smoking cessation. (*A*) Upon smoking cessation, genes upregulated by current smoking in the nasal epithelium are downregulated for 24 weeks. (*B*) Smoking cessation causes genes that are downregulated by current smoking in the nasal epithelium to increase in expression up to 16 weeks. Data are presented as mean ± SEM and analyzed using Student's parametric *t* test. Statistical significance to baseline (Week 0) is indicated by *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$. $n = 8$.
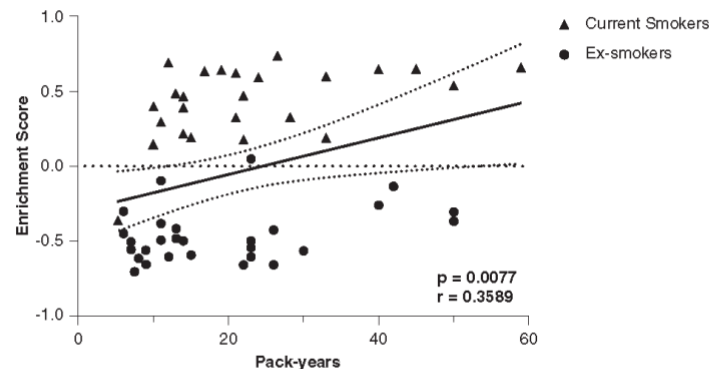
Figure 4. Correlation between the gene variation score of smoking-related upregulated genes and pack-years. A gene enrichment score of 0 is indicated by a dotted line, a solid regression line illustrates the direction of correlation between the *X* and *Y* variables, with 95% confidence intervals indicated by a dotted line. Triangles represent current smokers. Circles indicate ex-smokers. Data were analyzed by nonparametric Spearman's correlation (*r* = correlation coefficient). Statistical significance was determined at $P < 0.05$. $n = 26$ (current smokers) and 29 (ex-smokers) for all analyses.

Differential DNA Methylation Patterns Exist between Current and Ex-Smokers

To investigate the effect of current smoking on epigenetics (DNA methylation), we conducted a differential DNA methylation analysis. We identified 18,814 CpG sites altered by current smoking, with 10,697 sites with higher methylation and 8,117 sites with lower methylation (FDR < 0.05) in current smokers relative to ex-smokers (Figure 5A). Table E3 lists the top significantly differentially methylated CpG sites. Figure 5B highlights the most significant CpG sites that are associated with either increased (red) or decreased (blue) methylation (Δβ value) in current smokers compared with ex-smokers. Figure E5 illustrates this relationship with respect to log-fold change.

The cellular composition of samples can affect results from differential DNA methylation analysis (28). We completed a sensitivity analysis by including the cellular deconvolution data as covariates in the linear model. These results reported that 95% of significantly affected methylation sites were retained when cell composition was adjusted for, compared with the original analysis.

Changes in DNA Methylation Are Associated with Changes in the Expression of Smoking-related Genes

The *cis*-eQTM analysis aimed to investigate whether an association exists between significantly differentially expressed genes and methylated CpG sites in response to current smoking. A total of 171 *cis*-eQTM relationships were identified with key smoking-related genes (*ALDH3A1, AHRR, BPIFA2, CYP1A1, CYP1B1, CYP1B1-AS1,* and *SLC7A11*); the most significant *cis*-eQTM relationships per gene are listed in Table E4. Figures 6A–6G illustrate the relationship between gene expression (counts per million) and DNA methylation status for the most significant CpG sites identified by eQTM analysis. Figure 6A indicates that increased methylation of cg19949948 correlates with decreased expression of *ALDH3A1* (FDR = 5.74×10^{-04}). Conversely, Figure 6G shows that increased methylation of the CpG site cg13779050 correlates with increased transcription of *SLC7A11* (FDR = 3.81×10^{-02}).

Mapping of Smoking Affected Genes in an Independent *In Vitro* Dataset

To explore a possible explanation for how DNA methylation sites may influence the expression of the identified genes, we
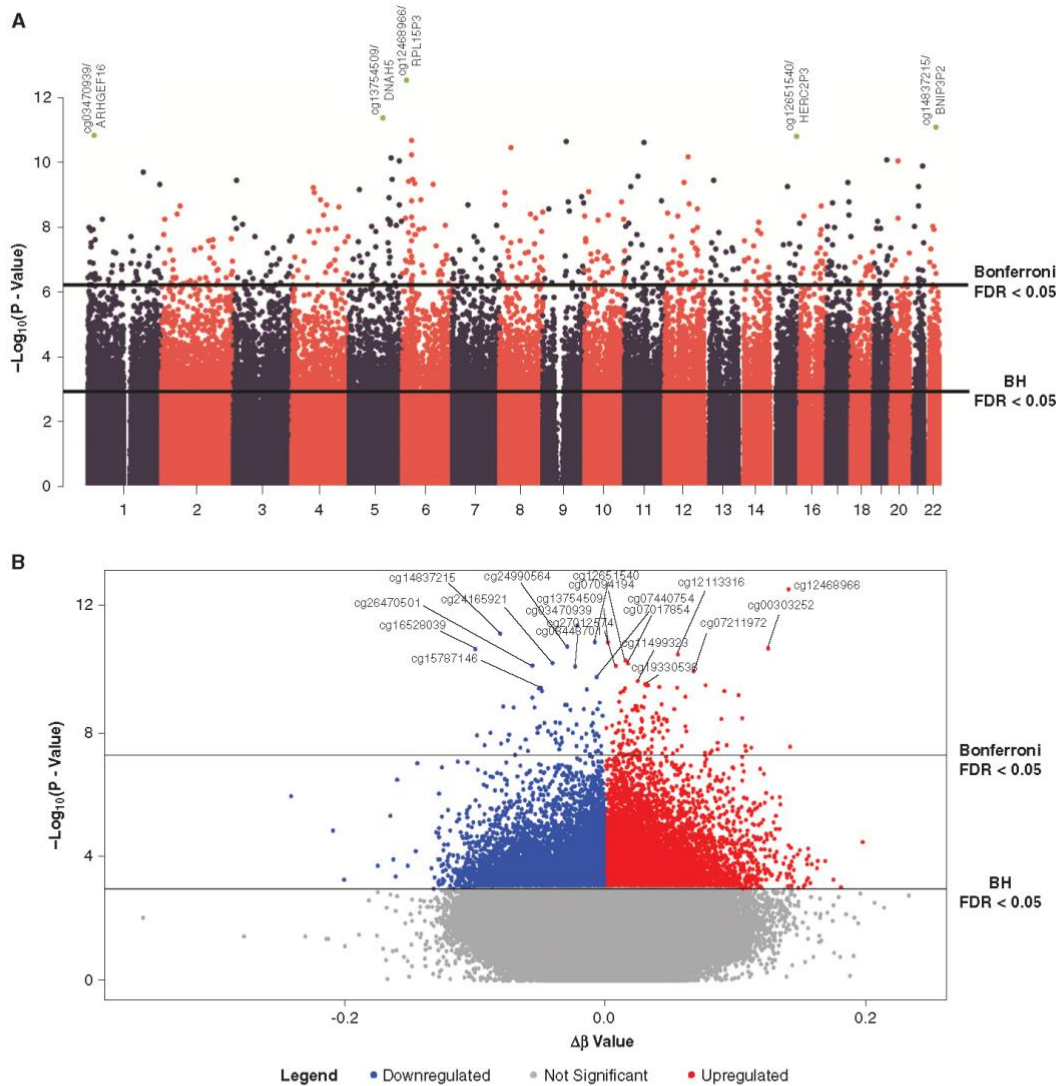
Figure 5. Manhattan and volcano plot visualization of lower and higher methylation of CpG sites in current smokers versus ex-smokers. (*A*) Manhattan plot showing the amount of differentially methylated sites on individual chromosomes, with the most significant sites highlighted and nearby genes indicated. (*B*) Volcano plot illustrating the CpG sites that are associated with either lower (blue) or higher (red) methylation and the $-\log_{10}$ (*P* value) plotted against the Δβ values for each CpG site. Genome-wide statistical significance is denoted in both figures by a solid black line at FDR-adjusted *P* value < 0.05 using the Bonferroni and BH method. *n* = 14 (current smokers) and 17 (ex-smokers) for all analyses. BH = Benjamin-Hochberg.

investigated publicly available ChIP-Seq datasets of well-known smoking-related transcription factors AHR, AHRR, and NRF2. With eQTM analysis, we identified three genes with a total of 17 CpG sites to be located in proximity to AHR and AHRR binding sites. No results were found for NRF2. Supplementary Figure E6 maps the transcript location for *CYP1B1*, *CYP1B1-AS1*, and *AHRR*, with AHR and AHRR transcription factor binding sites indicated in dark and light blue upstream of the gene. Multiple significant
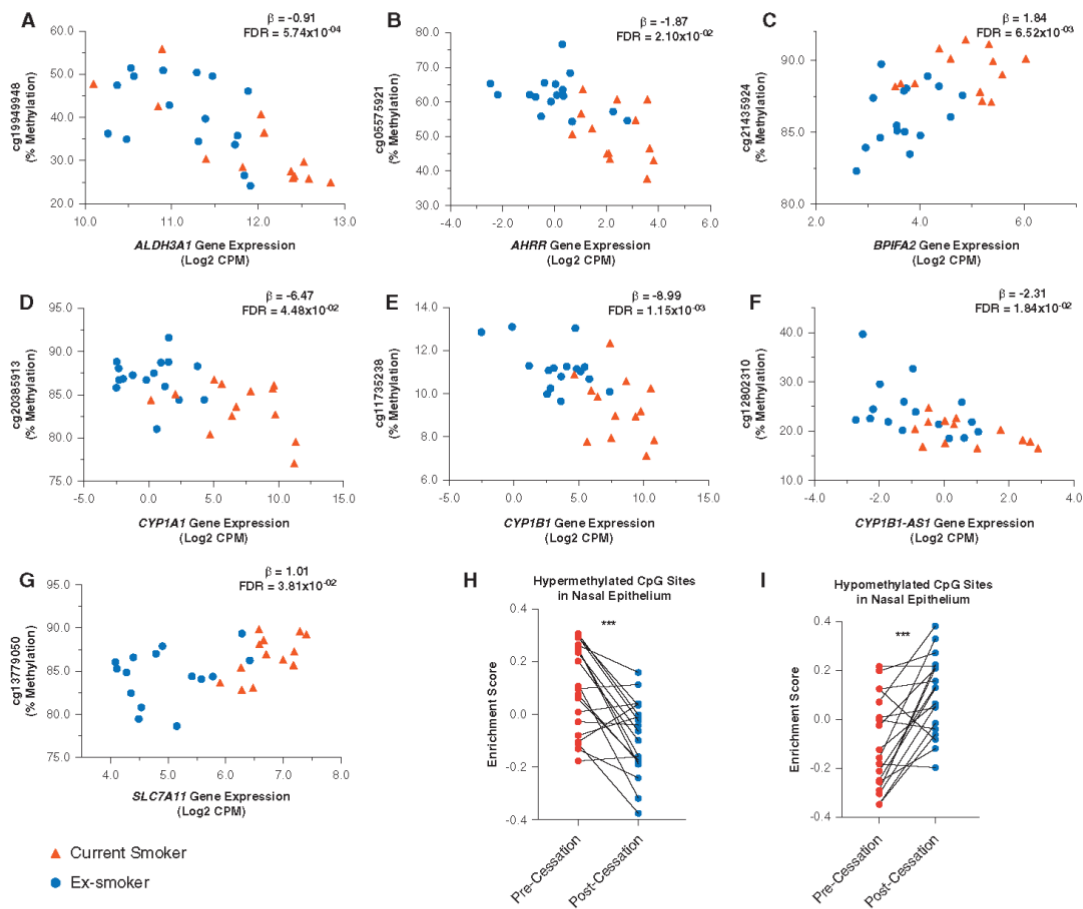
Figure 6. Representation of the relationship between DNA methylation status and gene expression in ex-smoker (blue) and current smoker (red) patients. (*A–G*) Percentage methylation values are reported on the *y*-axis and normalized gene expression is Log2 count per million (CPM) reported on the *x*-axis. The β-value indicates the slope of the correlation. All results returned an FDR <0.05. *n* = 13 (current smokers) and 16 (ex-smokers) for all analyses. (*H* and *I*) Gene set variation analysis data are represented as mean ± SEM and analyzed using a Student's paired parametric *t* test. Each dot represents an individual patient. Statistical significance is indicated by ***$P < 0.001$. *n* = 19. BPIFA2 = BPI fold-containing family A member 2; CYP1B1-AS1 = CYP1B1 antisense RNA 1; SLC7A11 = solute carrier family 7 member 11.

CpG sites identified by eQTM analysis were shown to be located within these sites, potentially influencing the binding of the transcription factors and therefore gene transcription.

We conducted a bootstrap analysis to increase the confidence that the identified list of CpG sites did not occur by chance. Across 1,000 permutations, an average of 1.23 CpG sites were identified with a 95% confidence interval of 1.15–1.31 CpG sites. Therefore, the total number of 17 CpG sites found in the original analysis was significant, with a *P* value of less than 0.0001. A frequency histogram has been included in the supplementary material to illustrate the results of the bootstrap analysis (Figure E7).

DNA Methylation of Specific CpG Sites Reverses with Smoking Cessation

Next, to investigate the long-term effect of smoking cessation on DNA methylation profiles and address the relevance for the lower airways, we used a bronchial biopsy dataset that followed current smokers for 1 year after smoking cessation (27). A total of 25 *cis*-eQTM CpG sites in the nasal epithelium were significantly altered in the reverse direction in bronchial biopsies after 1 year of smoking cessation (Table 3). GSVA of all differentially methylated CpG sites in the nasal epithelium indicated an overall reversal of DNA methylation status in the biopsies. Sites more

Table 3. List of *cis*-eQTM CpG Sites in Nasal Epithelium Altered after 1 Year of Smoking Cessation

		Nasal Epithelium Current Smoking			Bronchial Biopsy 1 Year after Smoke Cessation		
CpG Site	Gene	Log FC	FDR	Change in DNA Methylation	Log FC	FDR	Change in DNA Methylation
cg11183875	AKR1B10	−0.951	3.58×10^{04}	Decreased	0.341	2.55×10^{02}	Increased
cg04530175	AKR1B10	−1.112	4.52×10^{02}	Decreased	0.458	2.23×10^{02}	Increased
cg27638168	ALDH3A1	−0.694	3.15×10^{05}	Decreased	0.337	1.25×10^{02}	Increased
cg01375106	ALDH3A1	−0.375	4.15×10^{02}	Decreased	0.425	1.20×10^{02}	Increased
cg12828741	ALDH3A1	−0.518	2.30×10^{02}	Decreased	0.421	1.19×10^{02}	Increased
cg24752487	BPIFA2	−0.732	1.32×10^{02}	Decreased	0.424	4.16×10^{02}	Increased
cg02790305	BPIFA2	−0.720	4.72×10^{02}	Decreased	0.264	3.73×10^{02}	Increased
cg02162897	COL9A2	−1.701	1.36×10^{03}	Decreased	0.489	3.06×10^{02}	Increased
cg08761909	COL9A2	−1.086	8.77×10^{03}	Decreased	0.500	6.81×10^{03}	Increased
cg04968473	CYP1B1	0.565	1.99×10^{02}	Increased	−0.269	2.89×10^{02}	Decreased
cg01838114	CYP1B1	−0.887	8.92×10^{05}	Decreased	0.274	4.81×10^{02}	Increased
cg20408276	CYP1B1	−1.390	5.47×10^{03}	Decreased	0.432	2.65×10^{02}	Increased
cg26144569	CYP1B1	−0.663	4.73×10^{03}	Decreased	0.384	4.93×10^{02}	Increased
cg24577031	CYP1B1	0.435	3.97×10^{02}	Increased	−0.308	1.17×10^{02}	Decreased
cg12099669	CYP1B1-AS1	1.095	1.54×10^{02}	Increased	−0.543	2.20×10^{02}	Decreased
cg00305453	LHX6	−0.657	1.85×10^{02}	Decreased	0.274	4.62×10^{02}	Increased
cg13921444	MTNR1A	−0.734	2.09×10^{02}	Decreased	0.333	2.19×10^{02}	Increased
cg20385913	MTNR1A	−0.507	1.15×10^{03}	Decreased	0.442	2.96×10^{02}	Increased
cg13127741	SAA2	−0.507	2.70×10^{03}	Decreased	0.246	4.87×10^{02}	Increased
cg03296424	SAA4	−0.672	1.78×10^{04}	Decreased	0.319	2.12×10^{02}	Increased
cg06589222	SNTG2	−0.648	1.88×10^{02}	Decreased	0.243	1.98×10^{02}	Increased
cg07921828	STC1	−1.105	3.56×10^{02}	Decreased	0.524	2.61×10^{02}	Increased
cg26267490	STC1	0.812	2.72×10^{02}	Increased	−0.281	3.75×10^{02}	Decreased
cg03041510	SUSD2	0.398	1.05×10^{02}	Increased	−0.463	1.56×10^{02}	Decreased
cg05575921	TRHDE	−0.719	2.29×10^{02}	Decreased	0.341	1.90×10^{02}	Increased

Definition of abbreviations: *cis*-eQTM = *cis*-expression quantitative trait methylation; FDR = false discovery rate (Benjamini-Hochberg method); LogFC = log-fold change.

highly methylated in current smokers become more lowly methylated with 1 year of smoking cessation, and the opposite effect occurs for lowly methylated sites (Figures 6H and 6I).

Discussion

The genes that most consistently associated with current smoking in patients with asthma include *ALDH3A1*, *CYP1A1*, *CYP1B1*, *CYP1B1-AS1*, and *AHRR*. These genes were identified in the nasal epithelium and validated in controlled *in vitro* smoke exposure using ALI and *in vivo* airway cell models, demonstrating immediate changes in gene expression upon smoke exposure. We observe that genes that are increased by smoke exposure are downregulated with smoking cessation in upper and lower airways. Gene expression increased by current smoking correlate with peripheral blood immune cell counts, which may suggest a relationship between these genes and the innate immune response. Differentially methylated CpG sites were located in the proximity of key transcription factor binding sites, possibly

affecting the expression of the nearby genes. We also demonstrate changes in DNA methylation in the nasal epithelium are reflected in bronchi, highlighting a "field of injury" throughout the airways.

Genes significantly upregulated by current smoking in nasal epithelium of patients with asthma include *ALDH3A1*, *CYP1A1*, *CYP1B1*, and *AHRR*. *CYP1A1/B1* catalyze the breakdown of drug components producing reactive metabolites that can cause DNA damage (11, 19). Hence, prolonged expression of these genes increases the susceptibility of cancer development (29). *ALDH3A1* increases expression in response to aldehydes in smoke and other teratogenic components (30). Prolonged exposure to tobacco smoke products causes increased expression of these genes. This is reinforced by changes to DNA methylation, possibly at transcription factor binding sites, potentially influencing gene expression. The majority of the CpG sites indicated lower methylation levels in the nasal epithelium of current smokers, thus indicating that smoking may demethylate transcription factor binding sites, facilitating gene expression. Transcription

factor binding is known to be impacted by DNA methylation; however, the findings presented in our study require experimental validation to confirm this interaction. DNA methylation at AHRR and AHR binding sites need to be investigated to completely understand the dynamics between methylation and the ability of transcription factors to initiate gene transcription. AHRR has been linked to cigarette smoke and immune response in various tissues (1, 31–33). AHRR inhibits the transcription factor AHR, which regulates the expression of *CYP1A1* and *CYP1B1*. Dysregulated gene expression may prolong inflammation, leading to exacerbations in diseases such as asthma and COPD (34). *CYP1B1-AS1* is an antisense RNA that showed increased expression because of current smoking, indicating an effect of tobacco smoke on the regulation of protein translation.

Concurrent investigation of DNA methylation status and gene expression demonstrated a correlation between smoking status and methylation of CpG sites in proximity to smoking-related genes. Tobacco smoke has a heterogeneous effect on DNA

78

methylation status, potentially increasing or decreasing gene expression depending on the location of the site within a gene body (35). A similar investigation in small airway epithelium found changes in the methylation status of CpG sites associated with *ALDH3A1*, *CYP1A1*, and *CYP1B1* (18). The lower methylation of CpG site cg05575921 is observed to downregulate expression of *AHRR* in various tissues and cell types (19, 31, 36–38). Analysis of differential DNA methylation data in conjunction with independent ChIP-Seq data of AHR and AHRR illustrates the proximity of smoking-related DNA methylation sites to gene promoter regions. This reinforces the significance of these transcription factors as key regulators of the physiological response to tobacco smoke.

A correlation between the GSVA smoking score in the nose and the immune response is observed through increased peripheral blood neutrophil, lymphocyte, and monocyte absolute counts, whereas no association with peripheral eosinophil numbers was observed. Our findings agree with those of previous studies showing that current smoking increases peripheral blood neutrophils and that the nasal epithelium can reflect smoking-induced neutrophilic inflammation (39). Therefore, altered nasal gene expression and DNA methylation due to current smoking might relate to increased immune cell count and worsening asthma symptoms.

We identify a correlation between smoking-induced changes in DNA methylation and gene expression in nasal brushed cells. Investigating an independent *in vivo* smoking dataset demonstrated a reversal of smoking-associated upregulated genes within 4 weeks of cessation. A reversal in

the expression of smoking affected genes is sustained for 6 months with continued smoking cessation. In bronchial samples, we observed a reversal of DNA methylation status after 1 year of smoking cessation. We observed no strong cell-type contribution through cellular deconvolution. Surprisingly, a lack of goblet cell hyperplasia, which has been associated with smoking, was reported. This is possibly due to the maximal collection of these cells, reducing the sensitivity of the comparison. A consistent pattern and alteration of DNA methylation from the nose to bronchus reflects the concept of a "field of injury" throughout the respiratory tract caused by tobacco smoke exposure. Studies in adipose tissue support this, with cessation causing some reversal in methylation of CpG sites (19). Our analysis involved comparing nasal with bronchial epithelium CpG sites in different patient cohorts. Therefore, further investigation into the dynamics of changes in DNA methylation and the effect on gene expression within the same tissue in response to smoke is necessary.

A strength of our investigation is the ability to analyze both gene expression and DNA methylation in nasal epithelial samples from the same patients. This enables a direct study of the correlation between expression and DNA methylation status. However, there were a number of limitations in the current study. Our investigation compares data collected from varying cell types, limiting the ability to interpret relationships and correlations and to determine conclusions. However, these analyses, in combination with the primary analysis of paired gene and DNA methylation samples, offer valuable insight into potential mechanisms driving the biological response to tobacco smoke. We acknowledge the relatively low number of

patients in our study. Therefore, follow-up studies are required using a larger cohort, and the inclusion of a third never-smoker group in such a study would also be advisable, as it offers a more holistic understanding of the effect of tobacco smoke on the studied parameters.

This study confirms that current cigarette smoking alters epigenetic patterns and affects gene expression in nasal brushed cells of individuals with asthma. Current smoking affects the genes *ALDH3A1*, *CYP1A1*, *CYP1B1*, and *AHRR*, which are involved in physiological responses in detoxification and oxidative stress. Moreover, differentially methylated CpG sites due to current smoking demonstrate a relationship with the expression of these identified genes. Smoking cessation for 6 months indicates a reversal in gene expression pattern caused by smoking, with a similar pattern seen for DNA methylation levels after 1 year of cessation. Hence, a complex and dynamic relationship between DNA methylation and gene expression exists. Our findings contribute to the application and understanding of nasal gene expressions, such as the identification of appropriate biomarkers for disease prognosis and the elucidation of mechanisms contributing to pathological processes in smoking patients with asthma. ■

References

1. Kodal JB, Kobylecki CJ, Vedel-Krogh S, Nordestgaard BG, Bojesen SE. *AHRR* hypomethylation, lung function, lung function decline and respiratory symptoms. *Eur Respir J* 2018;51:1701512.
2. Jiménez-Ruiz CA, Andreas S, Lewis KE, Tonnesen P, van Schayck CP, Hajek P, *et al.* Statement on smoking cessation in COPD and other pulmonary diseases and in smokers with comorbidities who find it difficult to quit. *Eur Respir J* 2015;46:61–79.
3. Global Initiative for Asthma. Global strategy for asthma management and prevention online appendix 2020. 2020 [accessed 2021 Aug 9]. Available from: https://webmed.irkutsk.ru/doc/pdf/ginareport.pdf.
4. Bakakos P, Kostikas K, Loukides S. Smoking asthma phenotype: diagnostic and management challenges. *Curr Opin Pulm Med* 2016;22:53–58.
5. Dijkstra A, Vonk JM, Jongepier H, Koppelman GH, Schouten JP, ten Hacken NH, *et al.* Lung function decline in asthma: association with inhaled corticosteroids, smoking and sex. *Thorax* 2006;61:105–110.
6. Billatos E, Faiz A, Gesthalter Y, LeClerc A, Alekseyev YO, Xiao X, *et al.* Impact of acute exposure to cigarette smoke on airway gene expression. *Physiol Genomics* 2018;50:705–713.
7. Imkamp K, Berg M, Vermeulen CJ, Heijink IH, Guryev V, Kerstjens HA, *et al.* Nasal epithelium as a proxy for bronchial epithelium for smoking-induced gene expression and expression quantitative trait loci. *J Allergy Clin Immunol* 2018;142:314–317.e315.
8. Zhang X, Sebastiani P, Liu G, Schembri F, Zhang X, Dumas YM, *et al.* Similarities and differences between smoking-related gene expression in nasal and bronchial epithelium. *Physiol Genomics* 2010;41:1–8.
9. Sridhar S, Schembri F, Zeskind J, Shah V, Gustafson AM, Steiling K, *et al.* Smoking-induced gene expression changes in the bronchial airway are reflected in nasal and buccal epithelium. *BMC Genomics* 2008;9:259.
10. Poole A, Urbanek C, Eng C, Schageman J, Jacobson S, O'Connor BP, *et al.* Dissecting childhood asthma with nasal transcriptomics distinguishes subphenotypes of disease. *J Allergy Clin Immunol* 2014;133:670–678.e612.

11. Tekpli X, Zienolddiny S, Skaug V, Stangeland L, Haugen A, Mollerup S. DNA methylation of the CYP1A1 enhancer is associated with smoking-induced genetic alterations in human lung. *Int J Cancer* 2012;131:1509–1516.
12. Zakarya R, Adcock I, Oliver BG. Epigenetic impacts of maternal tobacco and e-vapour exposure on the offspring lung. *Clin Epigenetics* 2019;11:32.
13. Lee KW, Pausova Z. Cigarette smoking and DNA methylation. *Front Genet* 2013;4:132.
14. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet* 2011;12:529–541.
15. Qi C, Xu CJ, Koppelman GH. The role of epigenetics in the development of childhood asthma. *Expert Rev Clin Immunol* 2019;15:1287–1302.
16. Guida F, Sandanger TM, Castagné R, Campanella G, Polidoro S, Palli D, et al. Dynamics of smoking-induced genome-wide methylation changes with time since smoking cessation. *Hum Mol Genet* 2015;24:2349–2359.
17. Ringh MV, Hagemann-Jensen M, Needhamsen M, Kular L, Breeze CE, Sjöholm LK, et al. Tobacco smoking induces changes in true DNA methylation, hydroxymethylation and gene expression in bronchoalveolar lavage cells. *EBioMedicine* 2019;46:290–304.
18. Buro-Auriemma LJ, Salit J, Hackett NR, Walters MS, Strulovici-Barel Y, Staudt MR, et al. Cigarette smoking induces small airway epithelial epigenetic changes with corresponding modulation of gene expression. *Hum Mol Genet* 2013;22:4726–4738.
19. Tsai PC, Glastonbury CA, Eliot MN, Bollepalli S, Yet I, Castillo-Fernandez JE, et al. Smoking induces coordinated DNA methylation and gene expression changes in adipose tissue with consequences for metabolic health. *Clin Epigenetics* 2018;10:126.
20. Cox CA, Boudewijn IM, Vroegop SJ, Schokker S, Lexmond AJ, Frijlink HW, et al. Extrafine compared to non-extrafine particle inhaled corticosteroids in smokers and ex-smokers with asthma. *Respir Med* 2017;130:35–42.
21. Moses E, Wang T, Corbett S, Jackson GR, Drizik E, Perdomo C, et al. Molecular impact of electronic cigarette aerosol exposure in human bronchial epithelium. *Toxicol Sci* 2017;155:248–257.
22. Hijazi K, Malyszko B, Steiling K, Xiao X, Liu G, Alekseyev YO, et al. Tobacco-related alterations in airway gene expression are rapidly reversed within weeks following smoking-cessation. *Sci Rep* 2019;9:6978.
23. Procházková J, Strapáčová S, Svržková L, Andrysík Z, Hýžďalová M, Hrubá E, et al. Adaptive changes in global gene expression profile of lung carcinoma A549 cells acutely exposed to distinct types of AhR ligands. *Toxicol Lett* 2018;292:162–174.
24. Lehne B, Drong AW, Loh M, Zhang W, Scott WR, Tan ST, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biol* 2015;16:37.
25. Yang SY, Ahmed S, Satheesh SV, Matthews J. Genome-wide mapping and analysis of aryl hydrocarbon receptor (AHR)- and aryl hydrocarbon receptor repressor (AHRR)-binding sites in human breast cancer cells. *Arch Toxicol* 2018;92:225–240.
26. Wang X, Campbell MR, Lacher SE, Cho H-Y, Wan M, Crowl CL, et al. A polymorphic antioxidant response element links nrf2/smaf binding to enhanced mapt expression and reduced risk of parkinsonian disorders. *Cell Rep* 2016;15:830–842.
27. Willemse BW, ten Hacken NH, Rutgers B, Lesman-Leegte IG, Postma DS, Timens W. Effect of 1-year smoking cessation on airway inflammation in COPD and asymptomatic smokers. *Eur Respir J* 2005;26:835–845.
28. Houseman EA, Kim S, Kelsey KT, Wiencke JK. DNA methylation in whole blood: Uses and challenges. *Curr Environ Health Rep* 2015;2:145–154.
29. Port JL, Yamaguchi K, Du B, De Lorenzo M, Chang M, Heerdt PM, et al. Tobacco smoke induces CYP1B1 in the aerodigestive tract. *Carcinogenesis* 2004;25:2275–2281.
30. Jang JH, Bruse S, Liu Y, Duffy V, Zhang C, Oyamada N, et al. Aldehyde dehydrogenase 3A1 protects airway epithelial cells from cigarette smoke-induced DNA damage and cytotoxicity. *Free Radic Biol Med* 2014;68:80–86.
31. Zeilinger S, Kühnel B, Klopp N, Baurecht H, Kleinschmidt A, Gieger C, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One* 2013;8:e63812.
32. Gao X, Jia M, Zhang Y, Breitling LP, Brenner H. DNA methylation changes of whole blood cells in response to active smoking exposure in adults: a systematic review of DNA methylation studies. *Clin Epigenetics* 2015;7:113.
33. Arimilli S, Madahian B, Chen P, Marano K, Prasad GL. Gene expression profiles associated with cigarette smoking and moist snuff consumption. *BMC Genomics* 2017;18:156.
34. Stringer KA, Freed BM, Dunn JS, Sayers S, Gustafson DL, Flores SC. Particulate phase cigarette smoke increases MnSOD, NQO1, and CINC-1 in rat lungs. *Free Radic Biol Med* 2004;37:1527–1533.
35. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 2012;13:484–492.
36. Monick MM, Beach SR, Plume J, Sears R, Gerrard M, Brody GH, et al. Coordinated changes in AHRR methylation in lymphoblasts and pulmonary macrophages from smokers. *Am J Med Genet B Neuropsychiatr Genet* 2012;159B:141–151.
37. Bakulski KM, Dou J, Lin N, London SJ, Colacino JA. DNA methylation signature of smoking in lung cancer is enriched for exposure signatures in newborn and adult blood. *Sci Rep* 2019;9:4576.
38. Bojesen SE, Timpson N, Relton C, Davey Smith G, Nordestgaard BG. *AHRR* (cg05575921) hypomethylation marks smoking behaviour, morbidity and mortality. *Thorax* 2017;72:646–653.
39. Vachier I, Vignola AM, Chiappara G, Bruno A, Meziane H, Godard P, et al. Inflammatory features of nasal mucosa in smokers with and without COPD. *Thorax* 2004;59:303–307.

80

Chapter 5 – ZFX and ZFY differentially regulate hallmark features of asthma – inflammation, fibrosis and death

5.1 Introduction

Sexual dimorphism refers to the phenotypical differences between males and females, which extends beyond the sexual organs. Differences between the sexes exist at the physiological, molecular and genetic levels [24, 25]. The most fundamental differences between *Homo sapiens* males and females are the sex chromosomes, with human females carrying two X-chromosomes and males carrying one X-chromosome and one Y-chromosome. Therefore, an imbalance of gene dosage exists between the sex chromosomes, presenting as a possible biological factor driving sex differences. The X-chromosome is approximately three times the size of the Y-chromosome; the latter is historically considered a genetic wasteland [26]. However, genetic variations in the Y-chromosome have recently been implicated in coronary heart disease, adaptive immunity and inflammation [27]. We have previously highlighted the breadth and complexity of sexual dimorphism in multiple organ systems and a range of diseases [28]. Despite well-known patterns of sex differences in diseases, there remains a lack of research into the contribution of sex chromosomes to this phenomenon.

Asthma is a prevalent disease demonstrating a clear pattern of sexual dimorphism, which is well documented [29, 30]. Asthma is a chronic respiratory disease characterised by airway inflammation and variable airflow obstruction. 350 million people have asthma, with the disease-causing approximately 450,000 deaths worldwide [31]. Young males report higher rates of asthma diagnosis compared to young females. However, at puberty, there is a distinct shift post-puberty to higher rates and greater severity in females. This pattern persists into adulthood, with adult females demonstrating increased hospitalisation rates and worse outcomes related to adult asthma [32]. These sex differences manifest in hallmark disease processes such as differences in the ratio of T-cell populations [33] contributing to altered inflammatory responses in asthma. Sexual dimorphism in airway remodelling is also prominent. Adult females demonstrate a downward trend for the ratio between FEV_1 (forced expiratory volume in one second) to the total lung capacity, indicating more fibrosis and remodelling changes than males [34]. Furthermore, young males have lower expiratory rates than females despite both sexes having similar lung volumes, highlighting a physiological disadvantage [35]. Several factors have been suggested to account for the differences between males and females with asthma, from the sex hormones (estrogen and testosterone) [29] to differences in rates of lung development (dysanapsis) [36]. Both of these factors significantly change around puberty, where the distinctive increase in the female incidence of asthma is apparent. Many studies in recent years highlight the contribution of the sex

chromosomes to sexual dimorphism in disease susceptibility [37-40]. Nonetheless, the primary factor driving asthma sexual dimorphism remains unclear.

X-chromosome inactivation (XCi) occurs when one X-chromosome in females becomes hypermethylated, stopping the expression of genes from that chromosome. As a result, females 'theoretically' express genes from only one X-chromosome to equate to male X-chromosome expression. However, a subset of genes escapes XCi. These genes have highly similar homologous pairs on the Y-chromosome. Importantly, these genes are evolutionarily conserved genome regulators such as transcription factors, ribosomal proteins and histone demethylases.

The gene pair ZFX and ZFY are a set of zinc finger proteins that are highly conserved in vertebrates. These two genes have 91% similarity at the protein level, causing studies to theorise the two being homologous in function [41, 42]. However, this has not been confirmed by functional studies. ZFX has been widely investigated and has been associated with the progression of a range of cancers correlating with malignancy grade [43] and the renewal of embryonic and haemopoietic stem cells [44, 45]. Multiple members of the zinc finger protein family have also been linked with allergic asthma, with some contributing to patterns of sex differences [46-48]. ZFX regulates the cell cycle, where its knock-down results in reduced growth potential for glioma cells *in vivo* [43] and in lung cancer cells [49]. ZFY is shown to be increased in non-smoking individuals with lung cancer [50]. There is evidence of ZFX and ZFY affecting disease processes such as cell migration, proliferation, and fibrosis in various diseases [51]. These pathological features are prominent in asthma and demonstrate a sexually dimorphic pattern, with females indicating worse airway remodelling and fibrosis compared to males [34, 52]. To date, there is a distinct lack of studies investigating the role of ZFX and ZFY in respiratory diseases other than lung cancer.

Here we aim to explore and define the contribution of ZFX and ZFY to the regulation of the hallmark features of asthma – inflammation, fibrosis, and regulation of cell death. We establish a positive correlation between ZFX gene expression and FEV_1 % predicted scores in asthma and sex-specific downregulation of both ZFX and ZFY in asthmatic males. Using knockout cell lines, we discover differential regulation of inflammatory cytokines (CXCL8 and IL6), cell proliferation, changes to the extracellular matrix production (regulation of tenascin-C), and cigarette smoke-induced cell death between ZFX and ZFY. We highlight that ZFY cannot compensate for the loss of ZFX expression. We follow up these phenotyping experiments with transcriptomic analysis and identify potential mechanisms and pathways by which ZFX and ZFY may function. Finally, we generate a signature of genes altered by ZFX and ZFY and explore how this signature changes within asthma, indicating an association with lung function

in males with asthma. We demonstrate an imbalance in the function of ZFX and ZFY in asthma-related disease processes and present ZFX and ZFY as potential candidates contributing to sexual dimorphism in asthma.

5.2 Methodology

5.2.1 Analysis of X chromosome inactivation escapee gene expression

An independent dataset available on the gene expression omnibus (GSE179277[53]) was used to explore the change in gene expression of X and Y chromosome-linked genes pre and post-puberty. This study included RNA-sequencing data from nasopharyngeal swabs of patients. The study included patients from a wide age range of 0 to 89 years, with a total cohort size of 238. However, for this analysis, we were only interested in pre and post-onset puberty changes. It is accepted that physiological changes begin in young males and females by 10 to 12 years old, and puberty has finished by approximately 17 – 19 years old [54]. Patients older than this age range do not lie within the scope of this analysis and have been removed from the analysis. The patient demographics are summarised in Table 5.1. The average $\log_2$ counts per million (CPM) expression for all X chromosome inactivation escapee genes with Y-chromosome counterparts [1] were determined for patients pre and post-puberty. This change in average expression was plotted and analysed using GraphPad prism 9.

Table 5.1: Summary of the patient subpopulation from GSE179277.

	Pre-puberty (< 10 years old)	Post-puberty (> 9 years old)
n	58	25
Male, n (%)	24 (41.3%)	18 (72%)
Age range, yrs	0 - 9	10 - 19

5.2.2 Analysis of *ZFX* and *ZFY* expression in healthy and asthmatic patients

The publicly available single-cell RNA-seq 'Human Lung Cell Atlas' [15] and 'Lung Cell Atlas' [16] were accessed to analyse the gene expression of *ZFX* and *ZFY* in different epithelial cell types in non-asthmatic and asthmatic patients. The Indurian dataset was also used to assess *ZFX* and *ZFY* expression in males and females. Patient demographics of all studies are located in section 3.16.

5.2.3 Generation of CRISPR Cas9 genome deletion cell lines

Three unique ZFX and ZFY knockout cell lines were established in A549 cells, as described in section 3.2. A schematic representation of the overall study design and analyses is included in the supplement as Supplementary Figure S5.1.

5.2.4 Western blot

ZFX and ZFY knockouts were confirmed by western blot analysis. Anti-human ZFX (L28B6, #5419S, Cell Signalling) and anti-human ZFY (#PA5-68440, Invitrogen) were used to detect and visualise protein bands. Anti-human GAPDH was a loading control (#MAB374, Merck-Millipore). Detailed methodology is located in section 3.19.

5.2.5 Cell culture and treatments

All generated cell lines were maintained in DMEM growth medium. Detailed descriptions of cell maintenance techniques and treatments are in section 3.1.

5.2.6 Measurement of CXCL8 and IL6 protein secretion

Wildtype, ZFX KO and ZFY KO cell-free supernatants were collected 24 hours post-stimulation with TNFα (10 ng/mL), IL-1β (10 ng/mL) and TGF-β1 (10 ng/mL). The concentration of CXCL8 and IL6 was quantified by ELISA and was completed as described in section 3.3.

5.2.7 RNA-sequencing

Whole-cell RNA extracts from wildtype, ZFX KO and ZFY KO cells were collected 6 hours post-TNFα (10 ng/mL) and 48 hours post-TGF-β1 (10 ng/mL) stimulation and processed as described in section 3.10 and 3.11.

5.2.8 Differential gene expression analysis

Differential gene expression analysis was completed using the *Dseq2* package in R. Detailed description of differential gene expression (DGE) analysis is located in section 3.12.

5.2.9 LC-MS/MS proteomics analysis

Proteomics analysis was completed on wildtype, ZFX KO and ZFY KO cells as described in section 3.17. Protein lysate samples were collected at 24 hours post-TNFα (10 ng/mL) and 72 hours post-TGF-β1 (10 ng/mL) stimulation

5.2.10 Transcription factor enrichment analysis

Transcription factor analysis was completed for wildtype, ZFX KO and ZFY KO samples at baseline and after 6 hours of TNFα (10 ng/mL) stimulation, as described in section 3.13.

5.2.11 Gene set variation analysis (GSVA)

GSVA was completed as previously described in section 3.14, using a publicly available dataset on the Gene Expression Omnibus. This study completed chromatin immunoprecipitation sequencing (ChIP-Seq) for ZFX and ZFY in the male 22Rv1 cell lines (GSE145160) [55]. Genes that were bound by ZFX only, ZFY only and by both were identified, and three separate gene lists were generated. GSVA analysis was completed to track how the expression of these gene lists varied in each knockout cell line.

This analysis was also completed using the protein abundance dataset generated by LC-MS/MS. This analysis was completed to compare the cell lines' transcriptome and proteome, by tracking the production of DEGs to their corresponding proteins in the proteomics analysis dataset.

GSVA was also completed on the bronchial biopsies collected from asthmatic and healthy control patients in the Indurian dataset to analyse the association between knockout cell line gene expression and clinical phenotypes.

5.2.12 Analysis of biological pathways enriched in knockout cell lines

The g:Profiler online tool investigated gene ontology pathways enriched in differential gene expression analyses between ZFX KOs, ZFY KOs and wildtype cells. Pathway analysis was completed as described in section 3.15.

5.2.13 Wound healing assay

The wound healing assay was completed on wildtype, ZFX KO and ZFY KO cells in growth medium over a 72 hr period, as described in section 3.4.

5.2.14 Cell proliferation

The cellular proliferation rate of wildtype, ZFX KO and ZFY KO cells was determined by manual cell counting after 96 hr incubation in growth medium. Detailed methods are described in section 3.9.

5.2.15 Cell adhesion to fibronectin

The analysis of fibronectin-mediated cellular adhesion of wildtype, ZFX KO and ZFY KO cells after one-hour incubation in DMEM supplemented with 1% (v/v) FBS. Detailed methodology is described in section 3.5.

5.2.16 Measurement of ECM protein

Altered deposition of fibronectin (FN1), tenascin-C (TNC) and collagen 4α1 (COL4A1) extracellular matrix proteins from wildtype, ZFX KO and ZFY KO cells was completed as described in section 3.8.

5.2.17 Cigarette smoke extract (CSE) generation and analysis cell death response

Marlboro Red standard cigarettes (Philip Morris) were used to prepare CSE as described in section 3.6.

5.2.18 Cell viability assay

The cellular viability of wildtype, ZFX KO and ZFY KO cells after 24 and 48 hr CSE exposure was assessed using the MTT assay described in section 3.7.

5.2.19 Protein quantitative trait loci validation

Gene expression correlation with protein abundance levels was analysed using protein quantitative trait loci (pQTL) for wildtype, ZFX KO and ZFY KO cells as described in section 3.18.

5.3 Results

5.3.1 ZFX and ZFY are differentially regulated in asthma

Sex differences in the incidence of asthma reverse at puberty. As such, we investigated how gene expression of X-chromosome escapee genes and their Y-chromosome counterparts changes before and after puberty using the online dataset GSE179277 [53]. Interestingly, average X-chromosome gene expression does not change after puberty for either sex, whilst the mean Y-chromosome gene expression increases in males post-puberty (Figure 5.1A & 5.1B). We next investigated the correlation between XY chromosome gene pair expression and asthma severity measured by FEV_1 % predicted values. All XY gene pairs expression were analysed in asthma patients, with *ZFX* the only gene to demonstrate a statistically significant positive correlation with the FEV_1 % predicted (Figure 5.1C). As such, there is an association between the *ZFX* expression level and asthma lung function.

To investigate differences between *ZFX* and *ZFY*, their relative gene expression at a single-cell level was analysed using a publicly available dataset [15]. Supplementary Figure S5.2 shows that *ZFX* and *ZFY* are evenly expressed in all cell types, as they are not over-expressed in a specific cell subtype. When investigating changes in particular cell types, *ZFX* and *ZFY* expression vary between control and asthmatic patients in a cell-specific manner (Figure 5.1D & 5.1E). ZFX gene expression in basal cells increases in asthmatics compared to healthy patients, whilst ZFY gene expression decreases. Analysis of gene expression in male and female asthma patients revealed reduced *ZFX* and *ZFY* decreases only in male patients compared to healthy counterparts (Figure 5.1F & 5.1G). These results highlight the expression of *ZFX* and *ZFY* is sensitive to disease status, and their functions vary depending on the cell type. To explore *ZFX* and *ZFY* expression throughout the respiratory tract, we analysed the cohort of bronchial biopsies (Figure 5.1H). The sum of *ZFX* and *ZFY* expression in males equated to the total expression of *ZFX* in females (double dosage) in both bronchial and nasal cells. However, for proper balance in functional output to persist between the sexes, ZFY's activity must match that of ZFX. If ZFY has reduced functionality, then males will be at a disadvantage. These analyses establish a dynamic and complex relationship between ZFX, ZFY and asthma. Due to the unequal expression of *ZFX* and *ZFY* between the sexes, it is necessary to explore whether these genes have the same regulatory functions and activity in modulating disease-relevant processes.
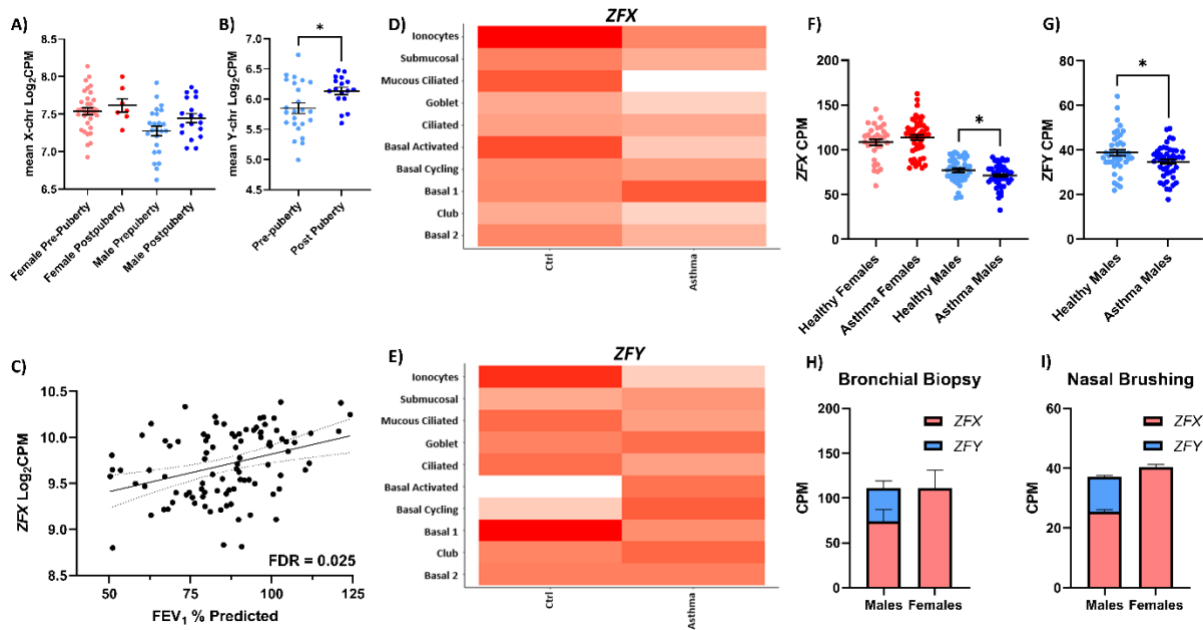
Figure 5.1: Analysis of *ZFX* and *ZFY* gene expression in asthma at a single cell level, bronchial biopsy, and nasal brushing. (A & B) Average expression of X-Y chromosome gene pair homologs pre-puberty (n = 34 females/24 males) and post-puberty (n = 7 females/ 18 males). (C) *ZFX* log$_2$ counts per million (CPM) gene expression plotted with FEV$_1$ % predicted value in patients with asthma, FDR-adjusted p-value determined using the Benjamini-Hochberg method; n = 96. Single-cell sequencing analysis of *ZFX* (D) and *ZFY* (E) expression across multiple epithelial cell types in both non-asthmatic control (n = 4 females/7 males) and asthmatic (n = 2 females/ 7 males) patients. Darker red = higher gene expression, white = no gene expression. *ZFX* (F) and *ZFY* (G) CPM gene expression in healthy (n = 33 females/42 males) and asthmatic (n = 48 females/46 males) patients. (H) Bronchial biopsy (n = 85 females/88 males); data presented as mean +/- SEM; pink colour = *ZFX* expression, blue colour = *ZFY* expression. (A – B & F – G) Data are presented as the arithmetic mean +/- SEM with unpaired parametric t-test, with statistical significance indicated by *p<0.05. *FDR = False discovery rate; FEV$_1$ % predicted = forced expiratory volume in one-second per cent predicted.*

5.3.2 ZFX and ZFY demonstrate different regulatory functions in inflammation and remodelling processes

To investigate the functional roles of ZFX and ZFY knockout, A549 cell lines were generated using CRISPR Cas-9. ZFX and ZFY knockout were confirmed using genome sequencing (Supplementary Figure S5.3) and western blotting (Figure 5.2A). *ZFX* and *ZFY* gene expression were analysed at baseline. *ZFX* gene expression increased in ZFX knockout cells only (Figure 5.2B), whilst *ZFY* gene expression was unchanged across all cell lines (Figure 5.2C). This indicates a feedback mechanism exists in ZFX-deficient cells, which is not apparent for ZFY.

ZFX and ZFY knockout cell lines were characterised for inflammation and fibrosis-related responses to determine if they have similar regulatory functions. CXCL8 and IL6 are important proinflammatory cytokines in asthma [56]. We observed ZFX knockout cells had suppressed TNFα-induced CXCL8 production compared to wildtype and ZFY knockout cells. In contrast, ZFX knockout cells produce elevated levels of IL6 compared to wildtype and ZFY knockout. CXCL8 and IL6 production from ZFY KO cells remained similar to wildtype cells (Figure 5.2D & 5.2E). Cells were also stimulated with the proinflammatory cytokines IL-1β and TGF-β1. The same patterns of CXCL8 and IL6 production were observed (Supplementary Figure S5.4). Despite CXCL8 and IL6 having highly similar functions and regulatory pathways, distinct opposing effects of ZFX KO were observed across multiple inflammatory stimuli. This indicates a complex regulatory role for ZFX in the immune response and suggests that ZFX and ZFY have distinctly different regulatory mechanisms.

Attachment, proliferation and wound healing are critical processes in remodelling the airways [52]. ZFX and ZFY knockout cell lines demonstrated opposite wound healing rates (Figure 5.2F). ZFY knockouts healed slower than wildtype cell lines, while ZFX knockouts healed significantly faster. Cell attachment to fibronectin was measured, where ZFX knockout cells demonstrated significantly increased attachment levels compared to wildtype and ZFY knockout cells (Figure 5.2G). ZFX knockout cells indicated a significantly slower proliferation rate than wildtype (Figure 5.2H), with a double time of 25 hours compared to 23.5 hours, respectively. ZFY knockouts reported a trend towards reduced proliferation (figure 5.2H), with a doubling time of 24.5 hours.

ZFY knockouts do not affect the inflammatory response but show a pattern towards regulating cell proliferation. Further, Figure 5.2I presents that ZFY knockout cells display increased survival in response to cigarette smoke extract (CSE) compared to wildtype cells. Continued survival by ZFY cells indicates a dysregulation of apoptotic processes in response to noxious stimuli.
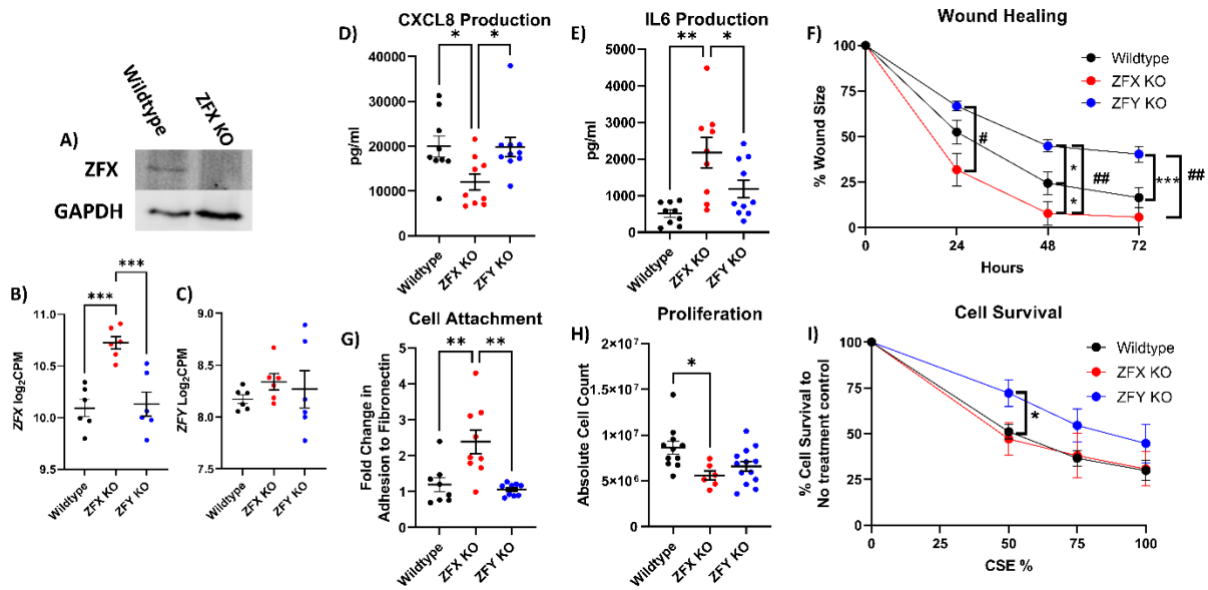
Figure 5.2: Characterisation of CRISPR-Cas9 generated ZFX (red) and ZFY (blue) knockout cell lines. (A) ZFX was confirmed by western blot. (B-E & G-H) the x-axis indicates the genotype of the cell line. Black dots represent wildtype cell lines, red represents ZFX knockout cells, and blue represents ZFY knockout cells. Log_2CPM expression of ZFX (B) and ZFY (C) across all cell lines. CXCL8 (D) and IL-6 (E) were measured after 24-hour TNFα (10 ng/mL) stimulation in cell-free supernatant by ELISA. (F) Cells were incubated at 37°C/5% CO_2 after a scratch wound was created in the growth medium. Wound closure was measured as a percentage compared to the initial wound size at zero hours. (G) Cell attachment was measured after one hour. (H) proliferation was measured after 96 hours of proliferation in a growth medium by cell counting. (I) Percentage cell survival compared to no treatment control was measured after 24-hour exposure to cigarette smoke extract (CSE) using the MTT assay. All data are presented as the arithmetic mean +/- SEM. One-way ANOVA statistical analysis with Tukey's correction for multiple comparisons used for B – E & G – H; *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001. Two-way ANOVA with two-stage step-up of Benjamini, Kreiger and Yekutieli correction for multiple comparisons method was used to analyse (F) and (I); *p-value < 0.05, *p-value < 0.001; # comparison ZFX KO vs ZFY KO. n = 6 – 10.

5.3.3 ZFY has similar but weaker activity compared to ZFX

We have now established that ZFX and ZFY have functional differences. To explore what genes and pathways are differentially regulated by ZFX and ZFY, we completed an RNA-Seq analysis. Cell lines were stimulated with TNFα for 6 hours. When ZFX is knocked-out, 3005 genes are differentially regulated compared to wildtype cells (Figure 5.3A), whilst 20 genes are differentially expressed when ZFY is knocked out (Figure 5.3B). Figures 5.3E and 5.3F indicate that genes differentially regulated in ZFX knockout cells are also dysregulated in ZFY knockout cells. GSVA was conducted to track the top 50 significant differentially expressed genes from each analysis within each knockout model to investigate this pattern further. We observe that genes upregulated in ZFX knockouts are less expressed in wildtype and ZFY knockout cells (Figure 5.3G). However, genes upregulated in ZFY knockout cells are upregulated in ZFX knockout cells (Figure 5.3H). Therefore, genes usually suppressed by ZFY are also suppressed by ZFX, but ZFY does not regulate most of the genes suppressed by ZFX. On the other hand, genes downregulated when both genes are knocked out are also downregulated in both ZFX and ZFY knockout cells, compared to wildtype (Figure 5.3I & 5.3J). This highlights that ZFX and ZFY share more common gene targets functioning as transcriptional activators rather than suppressors.
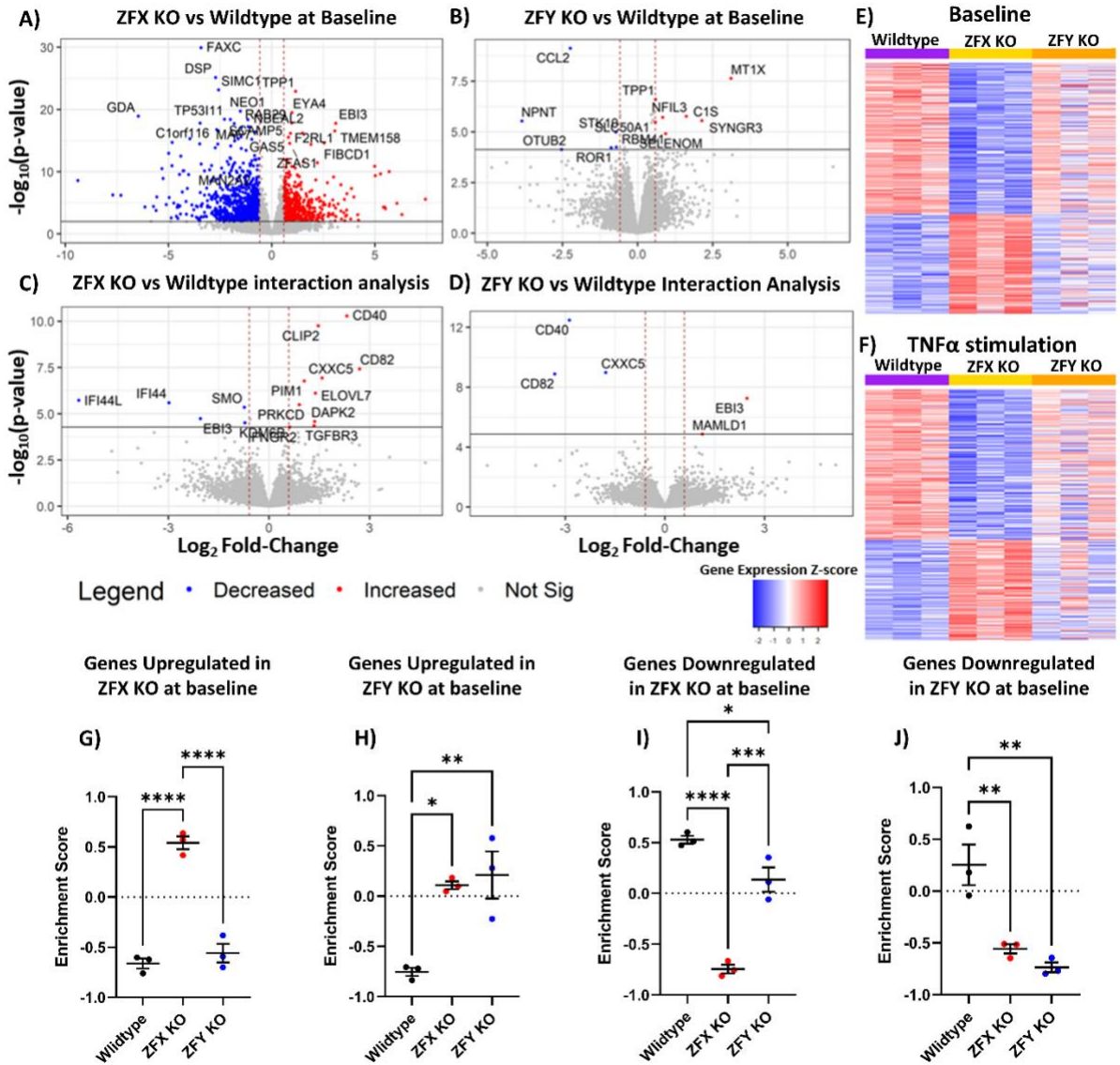
Figure 5.3: Differential gene expression and GSVA analysis of knockout ZFX and ZFY vs wildtype cell lines. (A – B) Volcano plot of -log₁₀(p-value) against log₂ fold-change in gene expression in both ZFX KO (A) and ZFY KO (B) against wildtype cells at baseline. (C – D) Volcano plot of differentially expressed genes (DEGs) when cell lines are TNFα-stimulated after 6 hours minus baseline DEGs. (C) ZFX KO against wildtype, and (D) shows ZFX KO against ZFY KO. (E – F) Supervised heatmap of differentially expressed genes in ZFX KO vs WT analyses at baseline (E) and TNFα stimulation (F). Increased genes are coloured red, and decreased genes are coloured blue. The genes are tracked for wildtype (purple), ZFX KOs (yellow) and ZFY KOs (orange). (G – J) GSVA of genes upregulated in ZFX KO (G) or ZFY KO at baseline (H) at baseline or downregulated (I – J). Data are presented as arithmetic mean +/- SEM and analysed using a one-way ANOVA with Tukey correction for multiple comparison testing. Statistical significance is indicated by *p<0.05, **p<0.01, ***p<0.001 and ****p<0.0001. n = 3 for all analyses.

5.3.4 ZFX and ZFY regulate a distinct set of genes

To further investigate the different regulatory functions of ZFX and ZFY, we compared the gene expression from ZFX knockout cells to ZFY knockout cells. However, to identify the unique differential gene signature between ZFX and ZFY, we removed significantly dysregulated genes common to each knockout cell line. As a result, we have identified a gene list that is differentially regulated between ZFX and ZFY. This provides greater insight into the distinct functions of both genes. We report 71 significantly upregulated (FC > 1) and 90 downregulated (FC < -1) differentially expressed genes between ZFX and ZFY knockout cells at baseline (Figure 5.4A). In Figure 5.4B, the heatmap highlights that these genes are distinctly dysregulated in these knockout cell lines compared to wildtype cells indicating the knockout cell lines, namely ZFY KO, differ significantly from wildtype cells. This is reinforced by GSVA analysis in Figure 5.4C – 5.4F, where wildtype cells generate an enrichment score between the ZFX and ZFY knockout cell lines whilst indicating statistical significance. This analysis reveals that the loss of ZFX does not solely drive the differentially regulated genes presented in Figure 5.4A, and ZFY has a functional output. If ZFY had a minimal function, then the enrichment score would be similar to what is produced by wildtype cells. Pathway analysis revealed cell migration, signalling and chemotaxis are enriched for ZFX knockout cells. In contrast, organism development and organisation pathways are increased in ZFY KOs (Supplementary Table S5.1). These pathways reflect the functional observations in Figure 5.2. ZFX knockout cells indicate increased production of IL6 (a key chemotactic cytokine) and increased cell migration. The pathways upregulated in ZFY KOs show a dysregulation of cell mitotic processes and structures, potentially contributing to reduced cell proliferation. Therefore, ZFX and ZFY have different and distinct regulatory pathways and mechanisms of action.
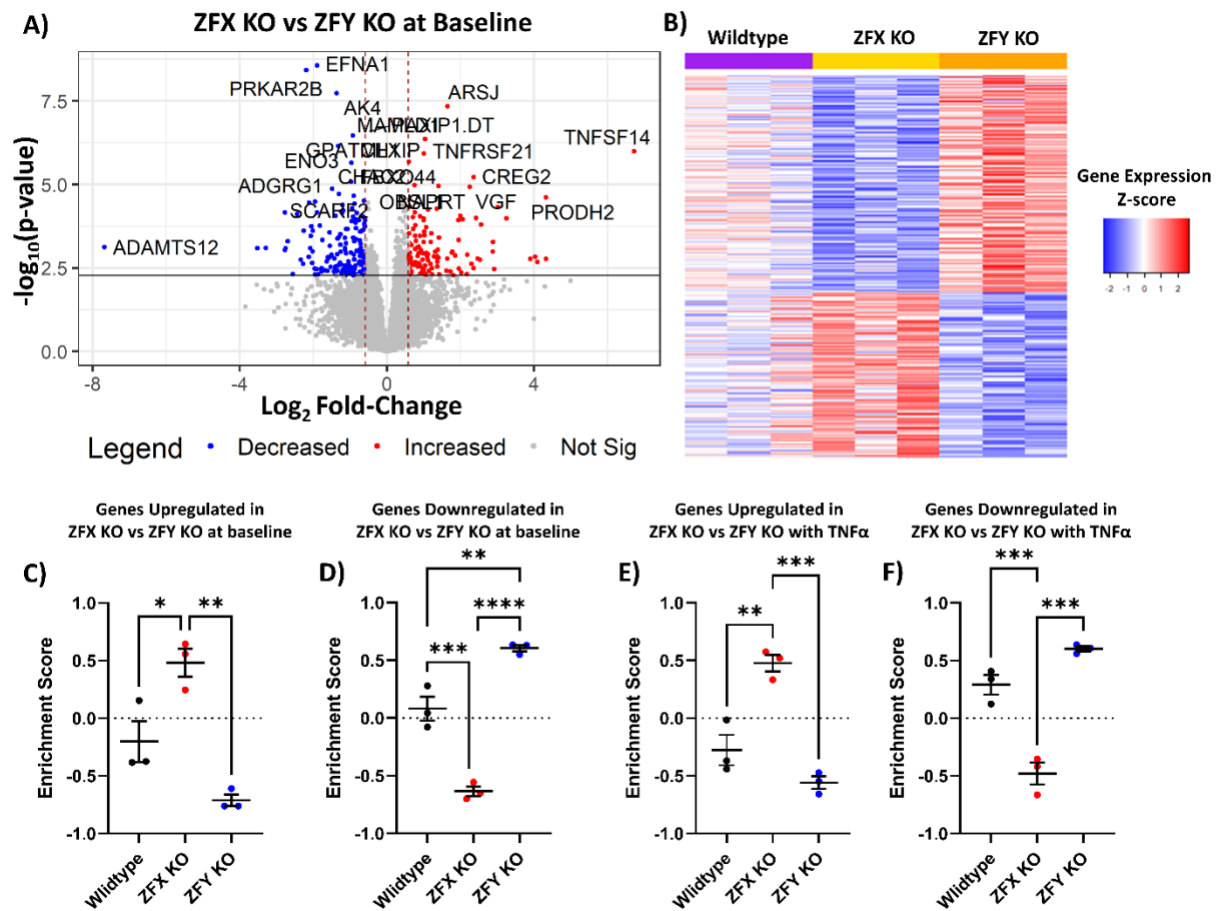
Figure 5.4: Analysis of a unique gene list dysregulated by ZFX and ZFY. (A) Volcano plot of log₂Fold-Change (x-axis) against -log₁₀(p-value) (y-axis) of differentially regulated genes between ZFX and ZFY knockout cell lines. Genes relatively upregulated in ZFX knockout cells are red, and genes relatively downregulated in ZFX are blue. Vertical red lines indicate a fold-change of |1.0|. Statistical significance was determined at FDR-adjusted p-value < 0.05 (indicated by the horizontal red lines) using the Benjamini-Hochberg method. (B) Supervised heatmap of differentially expressed genes in ZFX KO vs ZFY KO analysis at baseline. Increased genes are coloured red, and decreased genes are coloured blue. (C – F) GSVA of genes upregulated (C & E) or downregulated (D & F) in ZFX KO relative to ZFY KO at baseline (C & D) or after 6-hour TNFα (10 ng/mL) stimulation (E & F). Data are presented as arithmetic mean +/- SEM and analysed using a one-way ANOVA with Tukey correction for multiple comparisons. Statistical significance is indicated by *p<0.05, **p<0.01, ***p<0.001 and ****p<0.0001. n = 3 for all analyses.

5.3.5 ZFY is unable to compensate for ZFX deficiency

To further characterise and compare the function of ZFX and ZFY as transcription factors elements, we performed a GSVA using an independent ZFX knockout dataset GSE145160 [55]. ZFX was knocked out in male 22Rv1 cells, and RNA-sequencing was completed to identify differentially regulated genes. This data was used to quantify changes in gene expression in the knockout cell lines generated in this work by GSVA analysis. Genes up or downregulated were similarly altered in the ZFX knockout cells in the current study, with no significant effects seen in ZFY knockout cells (Supplementary Figure S5.5). This validates our findings and highlights ZFX's regulatory functions are consistent across different cell lines.

The previous study also completed ChIP-sequencing to identify the binding targets of ZFX and ZFY. We created a subset of these genes bound by ZFX, ZFY, or both. These were examined within our wildtype and knockout cell line datasets. Gene expression was investigated both at baseline and after 6-hour TNFα-stimulation. We observe that ZFX and ZFY bound genes are downregulated when ZFX is knocked out (Figure 5.5A & 5.5B). However, no change compared to wildtype is observed in ZFY knockout cells. As expected, genes regulated explicitly by ZFX only are down-regulated in ZFX knockout cell lines (Figure 5.5C & 5.5D), indicating ZFX functions primarily as a potent gene activator. Finally, genes bound by only ZFY are not altered in expression levels in ZFY knockout cells compared to wildtype, but these genes indicate a downregulation in ZFX knockout cells (Figure 5.5E & 5.5F). These results suggest redundancy in gene targets of both ZFX and ZFY. Figure 5.4 highlights the differential activity and function of ZFX and ZFY. When ZFX is knocked-out, ZFY cannot compensate for the deficiency in ZFX, whereas ok ZFY KOs, no change in gene expression is seen compared to wildtype cells. This indicates that ZFX can compensate for the lack of ZFY. Alternatively, the inability of ZFY to compensate for the loss of ZFX may be due to ZFY having significantly lower overall expression levels than ZFX (Figure 5.1H & 5.1I). Because ZFY expression does not increase when ZFX is absent (Figure 5.2B), there is a lack of abundance of ZFY to counteract the absence of ZFX. Nonetheless, this analysis highlights that some redundancy exists between ZFX and ZFY, and there is a distinct difference in the ability of ZFX and ZFY to compensate for the respective loss of the other.
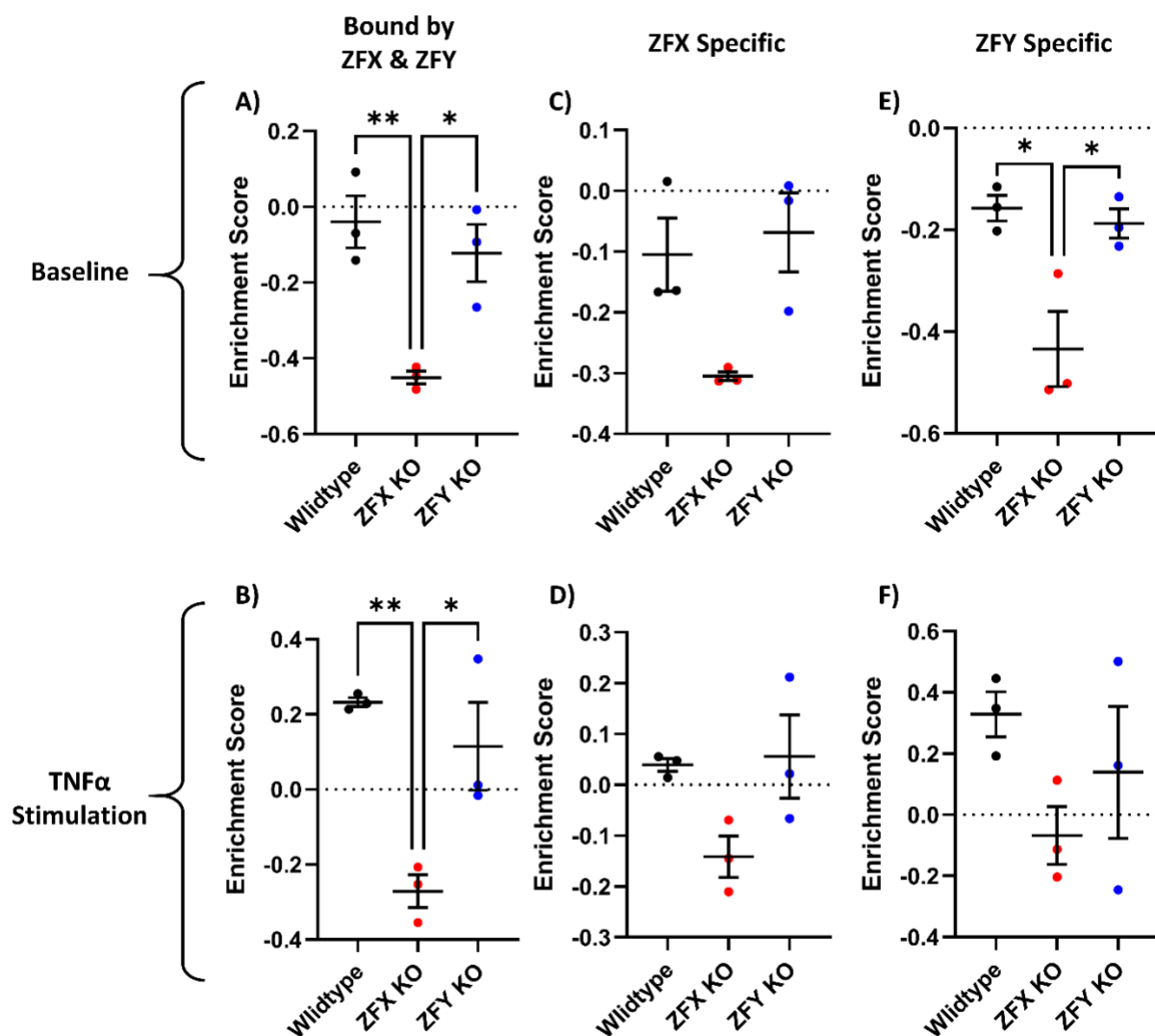
Figure 5.5: GSVA analysis of genes bound by ZFX and ZFY in wildtype and knockout cell lines. Genes bound by both ZFX and ZFY (A & B), ZFX only (C & D) and ZFY only (E & F) were tracked for expression changes at baseline (A, C & E) and after 6-hour TNFα stimulation (B, D & F). All gene lists were tracked in the RNA-seq gene expression dataset generated as part of this study. Data are presented as the arithmetic mean +/- SEM and analysed by one-way ANOVA with Tukey's correction for multiple comparisons. Statistical significance is indicated by *$p<0.05$, **$p<0.01$ and ***$p<0.001$; n = 3.

5.3.6 NF-κB activator subunits are differentially regulated by ZFX, affecting the proinflammatory response

CXCL8 and IL6 gene expression was analysed to investigate whether transcriptional or post-transcriptional factors account for differential protein production. Figures 5.6A & 5.6D demonstrate that when ZFX is knocked out, *CXCL8* expression is suppressed whilst IL6 expression is increased, as seen at the protein level (Figure 5.2A & 5.2B). Therefore, ZFX affects the transcriptional regulation of *CXCL8* and *IL6*. As such, we investigated the signalling cascade through transcription factor enrichment analysis to identify whether specific signalling protein activity is altered in the ZFX knockout model. We observed RELB is positively enriched, indicating increased activation in ZFX knockouts compared to wildtype and ZFY KO cells (Figure 5.6B). NKRF is also activated in ZFX knockout cells (Figure 5.6C). RELB activation is closely linked to NKRF production, associated with reduced CXCL8 production [57]. As such, the absence of ZFX promoted an increased activation of RELB and NKRF, which functioned to suppress CXCL8 production. Furthermore, we observed reduced activation of CEBPA but increased activation of CEBPB (Figure 5.6E & 5.6F). A positive correlation exists between transcription factor enrichment and gene expression (Supplementary Figure S5.6), where increased gene expression equates to increased activation of the respective pathway. However, in ZFX KOs at both baseline and post-TNFα stimulation, CEBPA expression and activation are reduced, indicating that loss of ZFX causes overall decreased production of CEBPA (Supplementary Figure S5.6C). Increased levels of CEBPA are reported to result in decreased IL6 production in a mouse model[58]. Therefore, dysregulation of CEBPA in ZFX knockout cells may increase IL6 production. Comparatively, CEBPB is well-known to complex with NF-kB proteins to activate IL6 gene expression synergistically [59]. Therefore, these findings indicate that ZFX regulates the RELB/NKRF mechanism to suppress CXCL8 and influences the activity of the CEBPA/CEBPB axis for the transcription of IL6.
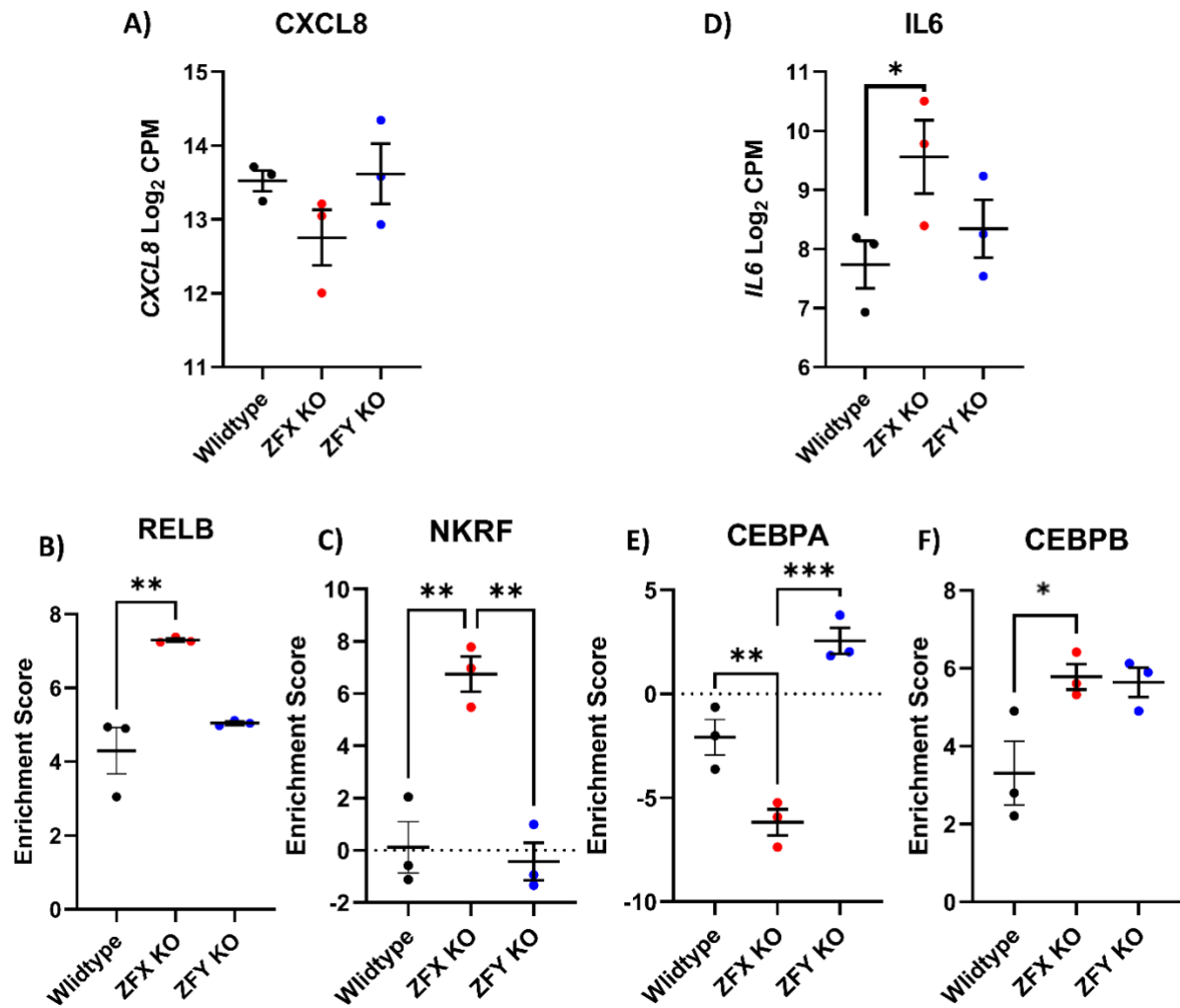
Figure 5.6: CXCL8 and IL6 gene expression and transcription factor enrichment analysis. (A & D) show $\log_2$CPM gene expression of CXCL8 and IL6 in the knockout and wildtype cell lines. (B, C, E & F) represent the enrichment score for activating different transcription factors after 6hr TNFα (10 ng/mL) stimulation. A positive enrichment score indicates the signalling pathway's activation, while a negative score indicates the pathway's suppression. All data are presented as the arithmetic mean +/- SEM and analysed by one-way ANOVA with Tukey's correction for multiple comparisons with statistical significance represented by $*p<0.05$, $**p<0.01$ and $***p<0.001$. n = 3 for all analyses.

5.3.7 Reduced ZFY activity compared to ZFX is also observed at a protein level

We recognised clear differential regulation of gene expression between ZFX and ZFY knockout cells at a transcriptomic level. Peptide-centric proteomic analysis confirmed these changes persisted at a protein level. Supplementary Figures S5.7A & S5.7B present a heat map of proteins differing in abundance in ZFX knockout cells compared to wildtype cells across all three cell lines. This indicates distinct differences between ZFX and ZFY knockout cell lines at a protein level. GSVA was completed to analyse whether significantly differentially expressed genes are similarly affected in the proteomics dataset. This analysis revealed that genes up or down-regulated in ZFX or ZFY knockout cell lines demonstrate the same direction of dysregulation at a protein level. For example, in Supplementary Figure S5.7D, for genes down-regulated in ZFX knockout cells, the corresponding proteins are also reduced in abundance. Our analysis confirms that although knocking out ZFY causes dysregulation of gene and protein expression, this is significantly less than ZFX, with ZFX affecting a unique set of proteins. It is important to note that there is a significant discrepancy in sensitivity between RNA-Seq and LC-MS/MS-based proteomics analysis, with many small proteins falling below the detection limit in this analysis. This notion may explain the lack of statistical significance observed in the GSVA comparisons of differentially expressed genes in ZFY knockout cells (Supplementary Figure S5.7G – S5.7I). Conversely, this observation may be caused by ZFX and ZFY affecting protein production through post-translational processes. Therefore, we analysed gene and protein expression, demonstrating a clear correlation between gene expression and protein level (Supplementary Figure S5.8). As such, this analysis confirms that the differential functions of ZFX and ZFY regulate gene expression, which can be observed with a direct change in the level of protein production.

5.3.8 ZFX and ZFY regulate cell adhesion and proliferation through integrins

Our initial characterisation study identified that cellular proliferation, adhesion, and wound healing were dysregulated in ZFX and ZFY knockout cell lines (Figure 5.2). These phenotypes contribute to fibrotic processes in disease, such as airway remodelling. Therefore, we conducted RNA-sequencing on cell lines at baseline and after 48-hour TGF-β1 stimulation. Differential gene expression analysis was performed between all cell lines. When ZFX is knocked out, 1442 genes are differentially expressed to wildtype cells at baseline (Figure 5.7A). In contrast, ZFY knockout cells displayed 115 differentially expressed genes at baseline (Figure 5.7C). 559 genes were differentially expressed between ZFX and ZFY knockout cell lines (Figure 5.7E). When stimulated with TGF-β1, knocking out ZFX caused significantly more differentially expressed genes than ZFY knockout cell lines. G: profiler analysed the molecular pathways enriched in the knockout cell lines (Table 5.2). This analysis revealed that in ZFX knockout cells, fibronectin binding, cell migration and integrin-mediated adhesion are positively enriched. ZFY knockout cells displayed reduced enrichment for mitotic centrosomes and proteins needed for microtubule organisation. As such, pathway analysis reveals that ZFX potentially mediates cell migration and adhesion processes, whilst ZFY affects cell proliferation and mitotic processes.
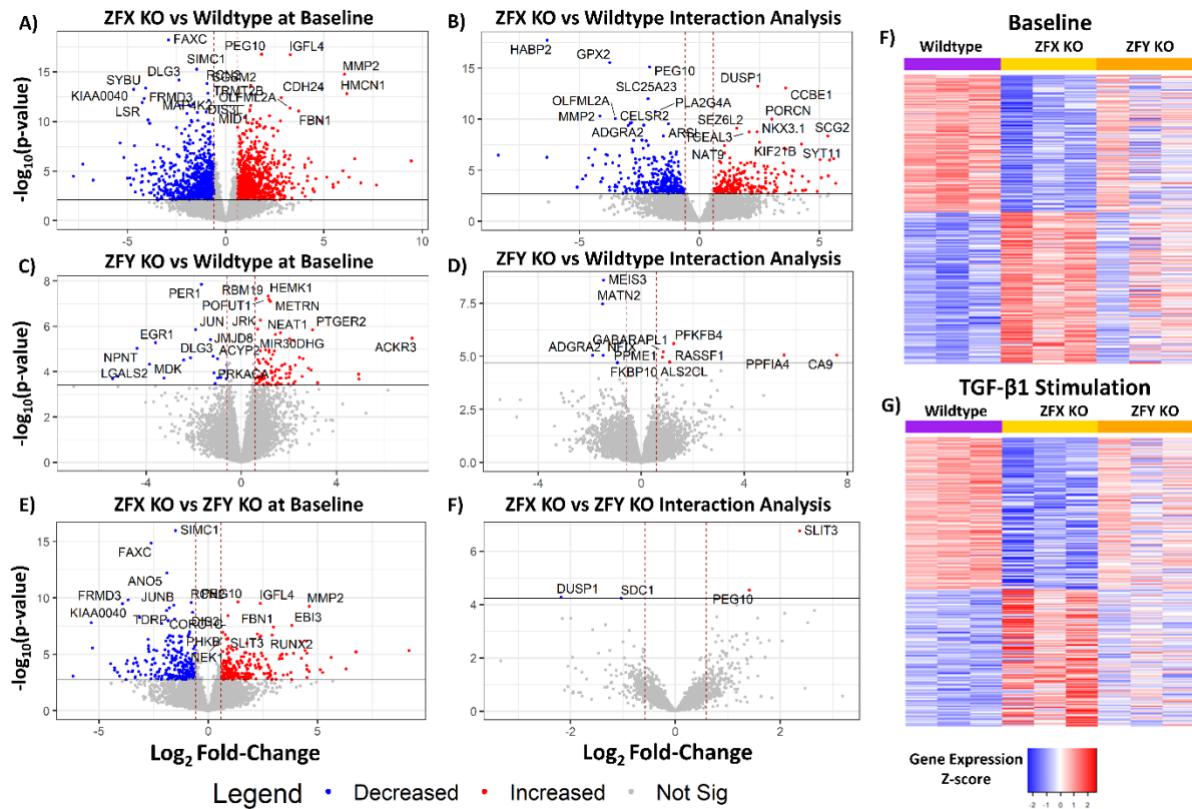
Figure 5.7: Volcano plot and heatmap visualisation of differentially expressed genes after 48-hour TGF-β1 stimulation. (A – F) Volcano plot of -log$_{10}$P-value (x-axis) against log$_2$fold-change (y-axis) in gene expression between all cell lines. (E – F) Delta gene expression analysis comparing DEGs after 48-hour TGF-β1 stimulation to baseline DEGs. Vertical red lines indicate a fold-change of |1.0|. Statistical significance was determined at FDR-adjusted p-value < 0.05 (indicated by the horizontal red lines) using the Benjamini-Hochberg method. (F – G) Supervised heatmap of differentially expressed genes in ZFX knockout cells compared to wildtype at baseline (F) and after 48-hour TGF-β1 stimulation (G). Genes are tracked for wildtype (purple), ZFX KO (yellow) and ZFY KO (orange). Increased genes are coloured red, and decreased genes are coloured blue. n = 3.

Table 5.2: Summary of g: Profiler pathway analysis using significantly upregulated genes in ZFX KO and ZFY KO cell lines compared to wildtype at baseline after 48 hours.

ZFX KO at Baseline		ZFY KO at Baseline	
Enriched Pathway	**FDR**	**Enriched Pathway**	**FDR**
Coreceptor Activity	8.12×10^{-4}	Loss of Nlp from mitotic centrosomes	2.30×10^{-2}
Fibronectin Binding	1.25×10^{-2}	Loss of proteins required for interphase microtubule organisation from the centrosome	2.30×10^{-2}
Positive regulation of cell migration	4.89×10^{-4}	Centrosome maturation	4.34×10^{-2}
Positive regulation of cell motility	7.12×10^{-4}	Recruitment of mitotic centrosome proteins and complexes	4.34×10^{-2}
Cell adhesion mediated by integrin	3.37×10^{-2}		

Definition of abbreviations: FDR = false discovery rate; Nlp = ninein-like protein

5.3.9 ZFX regulates changes to ECM proteins and remodelling processes

As key pathways of fibronectin-binding, cell migration and integrin-mediated adhesion are enriched, we investigated the expression of genes involved in these pathways. Fibronectin, tenascin-C and collagen 4α1 are closely associated with these pathways. Differential gene expression analysis revealed that all genes were increased in ZFX KOs compared to wildtype, with only collagen 4α1 showing a significant increase in ZFY KO cells (Supplementary Figure S5.9). As such, we validated these changes in gene expression at a protein level with ECM ELISAs. We confirmed that knocking out ZFX significantly increases fibronectin, tenascin C and collagen 4α1 compared to wildtype cells (Figure 5.8A – 5.8C). Fibronectin and TNC are both implicated in promoting increased adhesion and cell motility. Pathway analysis revealed integrin mediation of cell adhesion as a mechanism dysregulated in ZFX knockout cells. Integrin subunit β3 (*ITGB3*), Integrin subunit α4 (*ITGA4*) and integrin subunit β3 (*ITGB3*) were all dysregulated in ZFX knockout cells (figure 5.8D – 5.8F). High levels of *ITGB3/A4* and decreased *ITGB8* are linked with increased adhesion to fibronectin and increased cell migration. Therefore, ZFX mediates cell migration and attachment by producing the ECM proteins fibronectin and tenascin-C and their interaction with integrin receptors.
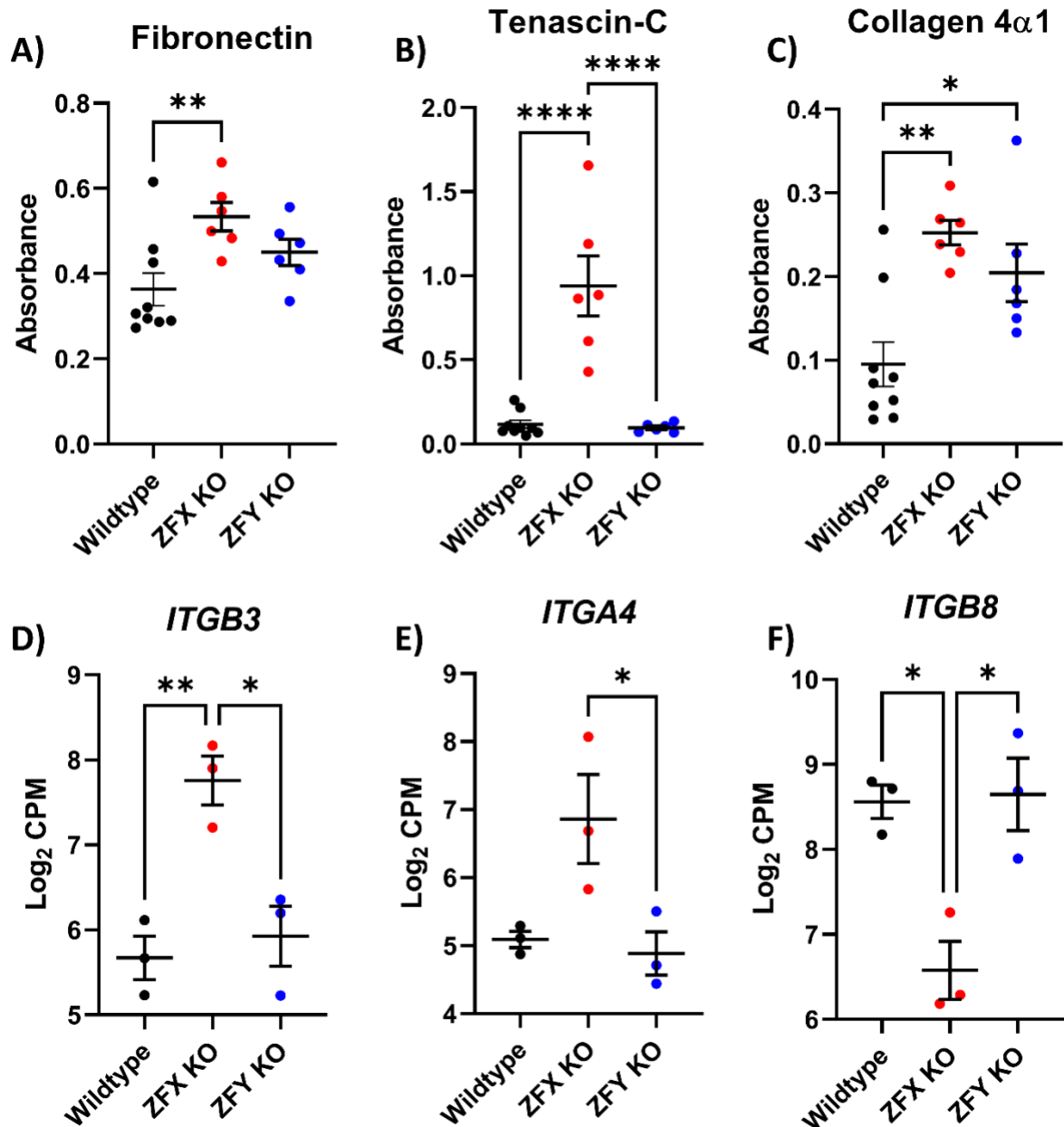
Figure 5.8: Differential regulation of ECM proteins and integrins by ZFX and ZFY. (A – C) ECM ELISA measured Fibronectin, Tenascin C and Collagen 4α1. Higher levels of absorbance (y-axis) indicate higher levels of the relevant protein; n = 6 – 9. (D – F) Log_2CPM gene expression of ITGB3, ITGA4 and ITGB8 measured by RNA-Seq. All data are presented as the arithmetic mean +/- SEM and analysed using one-way ANOVA with Tukey's correction for multiple comparisons. Statistical significance is indicated by *p<0.05, **p<0.01, ****p<0.0001; n = 3. *ITGB3 = Integrin-β3; ITGA4 = Integrin-α4; ITGB8 = Integrin-β8.*

5.3.10 Genes affected by ZFX are differentially regulated in asthma

To explore how genes regulated by ZFX and ZFY are altered in asthma, we completed a GSVA using significantly up or downregulated genes in the knockout cell lines at baseline. We tracked their expression in bronchial biopsies from healthy and asthmatic patients [18]. Genes upregulated in ZFX KOs are decreased in asthma (Figure 5.9A). Stratification by sex and asthma status revealed these genes are similarly reduced in both males and females with asthma (Figure 5.9B). GSVA scores were correlated with FEV_1 % predicted values for each patient using a multiple linear regression model correcting for cigarette smoking pack years, To understand whether these gene expression changes affect clinical outcomes. A significant positive correlation was reported in asthmatic males but not in females for genes upregulated in ZFX KOs (Figure 5.9C & 5.9D). As this effect was more substantial in males, we followed up this analysis using genes upregulated in ZFY KOs. Genes upregulated in ZFY KO cells demonstrated a positive relationship with FEV_1 % predicted in asthmatic males (Figure 5.9E). No effect was reported in females (Figure 5.9F). This data highlights that although similar gene pathways are altered between males and females with asthma, distinct physiological mechanisms contribute to sexually dimorphic health outcomes. Therefore, we identify that genes and pathways regulated by ZFX and ZFY contribute to clinical lung function outcomes in asthmatic males. These results highlight an active role for ZFX and ZFY in asthma, providing impetus for future studies.
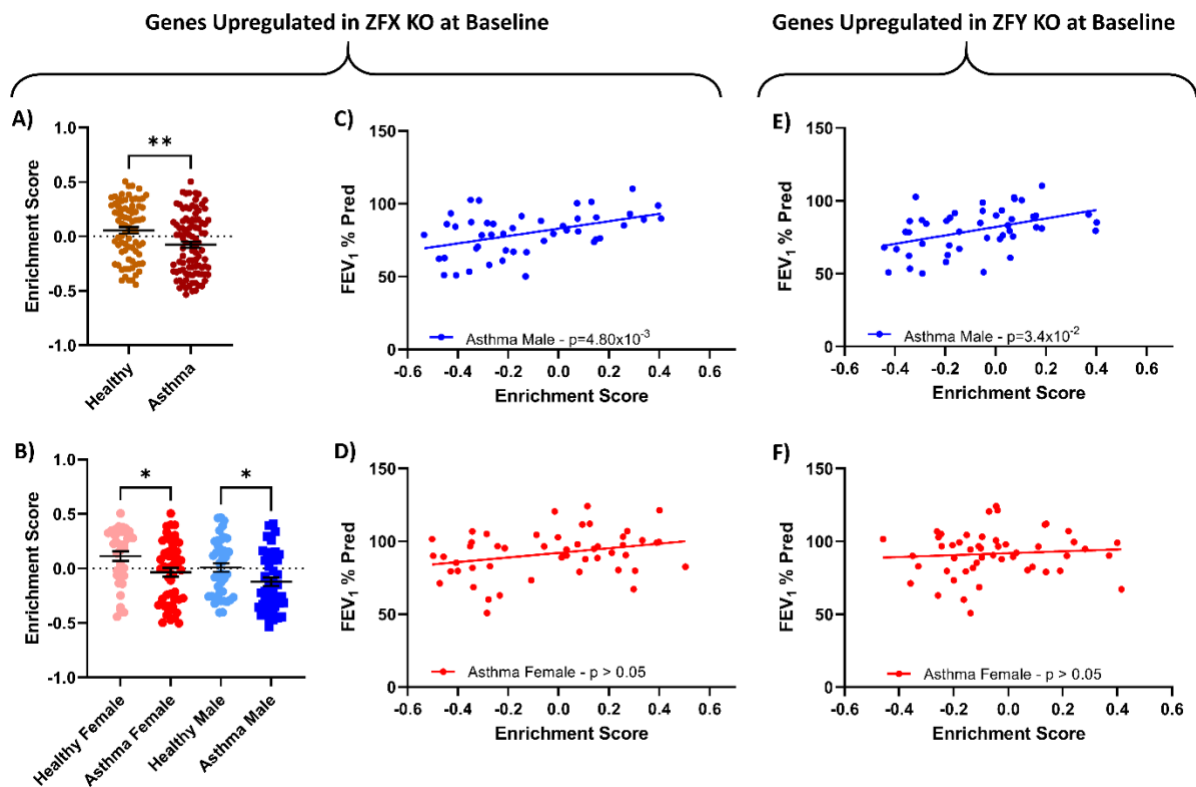
Figure 5.9: GSVA of genes upregulated in ZFX (A-D) and ZFY (E-F) knockout cells compared to wildtype in asthma patients. A subset of the 50 top significantly increased genes ZFX and ZFY knockout cells were analysed for how they change in patients with and without asthma. (A) Comparison of genes upregulated in ZFX KOs change in healthy (n=77) and asthma (n=96) patients, analysed by parametric t-test. (B) compares genes upregulated in ZFX KOs in healthy (n=35 females/ 42 males) and asthmatic (n=50 females/ 46 males) patients stratified by sex, analysed by one-way ANOVA with Tukey's correction for multiple comparisons. Data presented as the arithmetic mean +/- SEM. (C-F) The GSVA enrichment scores of genes upregulated in ZFX and ZFY KOs were analysed using a multiple linear regression model with Bonferroni correction for multiple comparisons to assess whether these scores correlate with FEV_1 % predicted measurements in asthma patients. Statistical significance was determined at $p<0.05$; n=50 females/ 46 males.

5.4 Discussion

This study investigated the function of ZFX and ZFY in the context of the hallmark features of asthma; inflammation, fibrosis and response to noxious stimuli. We show that ZFY and ZFY expression changes in asthma and their regulatory pathways correlate with lung function, potentially contributing to patient outcomes. We also demonstrate that ZFY cannot compensate for the lack of ZFX, whilst when ZFY is knocked out, cell death becomes aberrant. The characterisation of disease-relevant outputs paired with transcriptomic analysis enabled a multi-layered analysis exploring pathways and mechanisms causing the observed phenotypical differences between cell lines. ZFX and ZFY demonstrate distinct and differential functions regulating critical disease processes.

This is the first study to demonstrate that ZFX has an immunoregulatory capacity by affecting the expression of CXCL8 and IL6 cytokines. Despite CXCL8 and IL6 having highly similar functions driving the infiltration and activation of immune cells in asthma [60, 61], they have distinct regulatory pathways [62]. Therefore, the specific and opposite dysregulation in ZFX knockout studies reveals ZFX as a pivotal factor in the crossover in IL6 and CXCL8 regulatory pathways. Despite these proinflammatory cytokines having similar functions, ZFX appeared to regulate their production differentially. ZFX KOs demonstrated suppressed CXCL8 production whilst IL6 production increased compared to wildtype cells. This pattern existed at a transcriptomic level, indicating factors regulating CXCL8 gene expression are different from factors that regulate the expression of IL6. Analysis of transcription factor enrichment revealed increased activation of RELB and NKRF, leading to lower expression of CXCL8. A study by Ho et al. showed that forced expression of RelB promoted NKRF in A549 cells. This correlated with decreased production of CXCL8 [57]. As such, we indicate a function of ZFX in regulating this pathway.

In contrast to CXCL8, we observed a significant increase in IL6 production in ZFX KOs. Transcription factor analysis revealed that CEBPA activity is decreased in ZFX knockout cells while CEBPB activity increases. These isoforms of transcription factors have a functional role in initiating IL6 transcription. An *in vivo* mouse model demonstrated that over-expression of CEBPA attenuated the production of IL6 [58]. Comparatively, CEPBPB acts synergistically with NF-κB to activate the transcription of IL6 [59]. Therefore, although CEBPB also initiates CXCL8 production [59], the increased activation of RELB and NKRF suppresses its production in the ZFX knockout model. Notably, the effect of knocking out ZFX on the above-stated pathways and genes differed from the ZFY knockout model. This indicates an inability of ZFY to recover the functionality of ZFX in an inflammatory context. This potentially impacts the ability of asthmatic males to recruit neutrophils, as shown in our GSVA correlation. There was a positive correlation between genes upregulated when ZFX was knocked out to neutrophil

numbers, only in asthmatic males. This potentially indicates that the presences of two copies of ZFX in females enables a more robust inflammatory response, whilst this is lost in males.

We further investigate the difference in activity between ZFX and ZFY using ChIP-Seq data generated in the publicly available dataset (GEO145160) [55]. Genes bound by ZFX or ZFY and by both proteins were differentially expressed in ZFX knockout cells but were unchanged from the wildtype cell line in ZFY knockouts. Considering this pattern, we posit that ZFX has a distinctly greater functional capacity with significantly higher activity. However, we show that ZFY gene expression is significantly lower than ZFX. Therefore, the inability of ZFY to make up for a loss of ZFX may be due to an inadequate amount of ZFY protein to compensate effectively. Interestingly, *ZFX* gene expression is upregulated in ZFX KOs, indicating sensitivity to ZFX loss, highlighting its importance to normal cell functions, which is absent for ZFY. This discrepancy in gene regulation between ZFX and ZFY may relate to differences in their functional activity. Palmer et al. conducted one of the first studies to compare the ZFX and ZFY homologs, where they identified the complete open reading frame [63]. The authors identify differences between the proteins within the zinc finger region, which might affect specific base recognition. The acidic domains indicate 87% similarity, which may result in differences in activity between the proteins. The DNA sequence divergence in this region between ZFX and ZFY possibly contributes to the functional differences between ZFX and ZFY, identified for the first time in this study. We believe these subtle differences may drive the observed differences described in this study. However, more direct and specific analysis of these factors is needed to holistically understand its contribution to functional differences between ZFX and ZFY.

In particular, our analysis of fibrotic processes highlights that ZFX and ZFY regulate distinct pathways affecting proliferation and adhesion. Pathway analysis revealed that ZFX affects integrin-mediated cell adhesion, migration, and motility genes. In ZFY KOs, mitogenic pathways, such as the dysregulation of centrosome assembly, were enriched. The functional effect of these pathways is observed in the differential wound healing rates. Multiple studies have similarly shown a reduced expression of ZFX suppresses proliferation [43, 49, 64], supporting the result of this study. However, ZFX knockout cells demonstrated a dramatically faster rate of wound healing, whilst when ZFY was abolished, the scratch wound never reached closure. In alignment, we report increased adhesion to fibronectin by ZFX knockout cells, whilst no change is seen for ZFY. This is potentially driven by the upregulation of integrin molecules, which mediate the attachment to fibronectin. In particular, ITGA4, which is increased in ZFX knockouts, is linked with cell motility and migration, while ITGB3 is a driver for fibronectin assembly. Of particular interest, we report a significant increase in the production of tenascin-C (TNC) in ZFX knockout cells only, which is commonly downregulated

in adult tissues [65]. TNC interacts and functions in unison with fibronectin promoting increased cell motility [66], which we observe in our study with both TNC and FN production increased from ZFX knockout cell lines. However, in contrast, TNC has been shown to impede cell adhesion to fibronectin [67]. This apparent contradiction is explained by the design of our attachment assay, where Chiquet-Ehrismann et al. coated plates with FN1 and TNC, and we coated plates with FN1 only. Therefore, the competitive binding interaction described in their study does not apply to our model. Our results identify that the abolition of ZFX drives increased production of TNC and FN1, promoting increased cell motility and migration mediated through interaction with integrin surface molecules. TNC is significantly increased in the airways of people with asthma [68], with TNC also noted to have an immunomodulatory function inducing IL6 gene expression [68, 69]. Considering the strong association between TNC and ZFX, we propose that the role and activity of ZFX contribute to the progression of asthma.

Notably, the functional effect observed by knocking out ZFX was not reflected in the ZFY knockout model. Instead, TNC levels remained similar to wildtype cells, and a distinct inability for wound closure was recorded. In addition, the absence of ZFY increased survival in response to cigarette smoke extract. Dysregulation of cell death is likely related to a malformation of mitogenic structures, such as centrosomes, as identified by pathway analysis. A strong relationship between loss of the Y-chromosome and worse outcomes in cancer has been well established, with ZFY being specifically down-regulated [70]. Van den Berge et al. identified *ZFY* as downregulated in COPD [71]. ZFY presents as a critical cigarette smoke response gene located on the Y-chromosome as it is upregulated in smoking patients with cancer [72]. Although both ZFX and ZFY have been shown to have tumour suppressor capacity [73], a similar pattern of survival in ZFX knockout cells compared to wildtype indicates that ZFY has a critical function in response to noxious stimuli.

An important strength of this investigation was our use of multi-layered data, where we utilised laboratory assays for RNA-sequencing and proteomics. This allowed the functional characterisation of ZFX and ZFY knockout cell lines and the identification of potential pathways and mechanisms causing the observed changes. However, the proposed mechanisms are not validated and require future experiments to explore the action and function of ZFX and ZFY. Further, we only completed our model using A549 cells, and there is potential for ZFX and ZFY to have cell-specific effects, and their expression is temporally dependent. For example, ZFX is implicated in the differentiation of B-lymphocytes [74]. As such, there is the potential for unique cell-specific mechanisms and cell-specific differences in the function of ZFX and ZFY. The exploration of these mechanisms remained beyond the scope of the current study. We have demonstrated previously unreported functional

differences between ZFX and ZFY. These differences are related to critical pathological processes of inflammation, fibrosis and cell death. As such, we highlight important pathways by which the imbalance in ZFX and ZFY expression may contribute to sexual dimorphism in asthma. Our findings provide an impetus for future studies to understand the inequality in genetic factors contributing to sex differences in asthma. This increases the potential to identify targets for clinical interventions.

5.5 Conclusion

This study shows for first-time the functional differences between ZFX and ZFY. We further identify novel pathways regulated by both proteins related to immunoregulation and fibrosis. We highlight that ZFX indicates a higher activity level than ZFY and functions in regulating CXCL8, IL6 and TNC production. This contributes to pathologically relevant processes of inflammation and remodelling. As such, this imbalance in function between these genes may contribute to sexual dimorphic processes and outcomes observed in the disease. Our findings contribute to the growing literature recognising the need to understand the mechanisms driving sex differences in asthma and other prominent disorders. This provides future studies with greater scope to understand the function of sex chromosome-linked genes such as ZFX and ZFY and identify novel, more effective clinical interventions.

5.6 Supplementary Figures
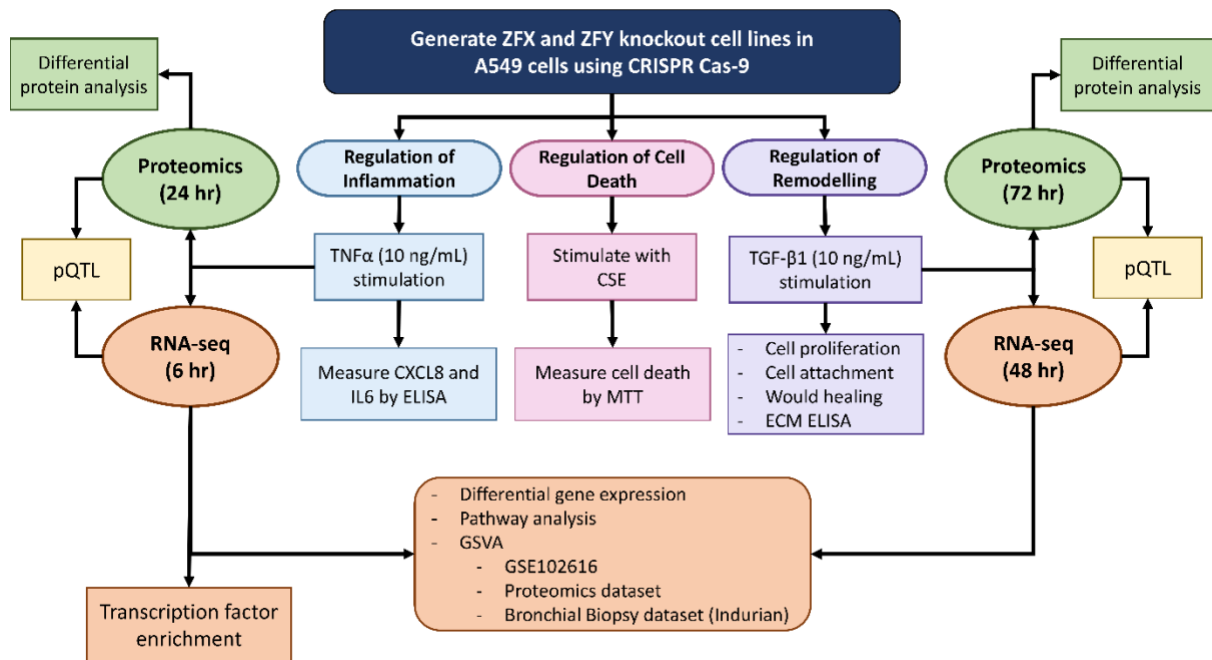
Figure S5.1: Schematic presentation of Chapter 5 study design and techniques used.
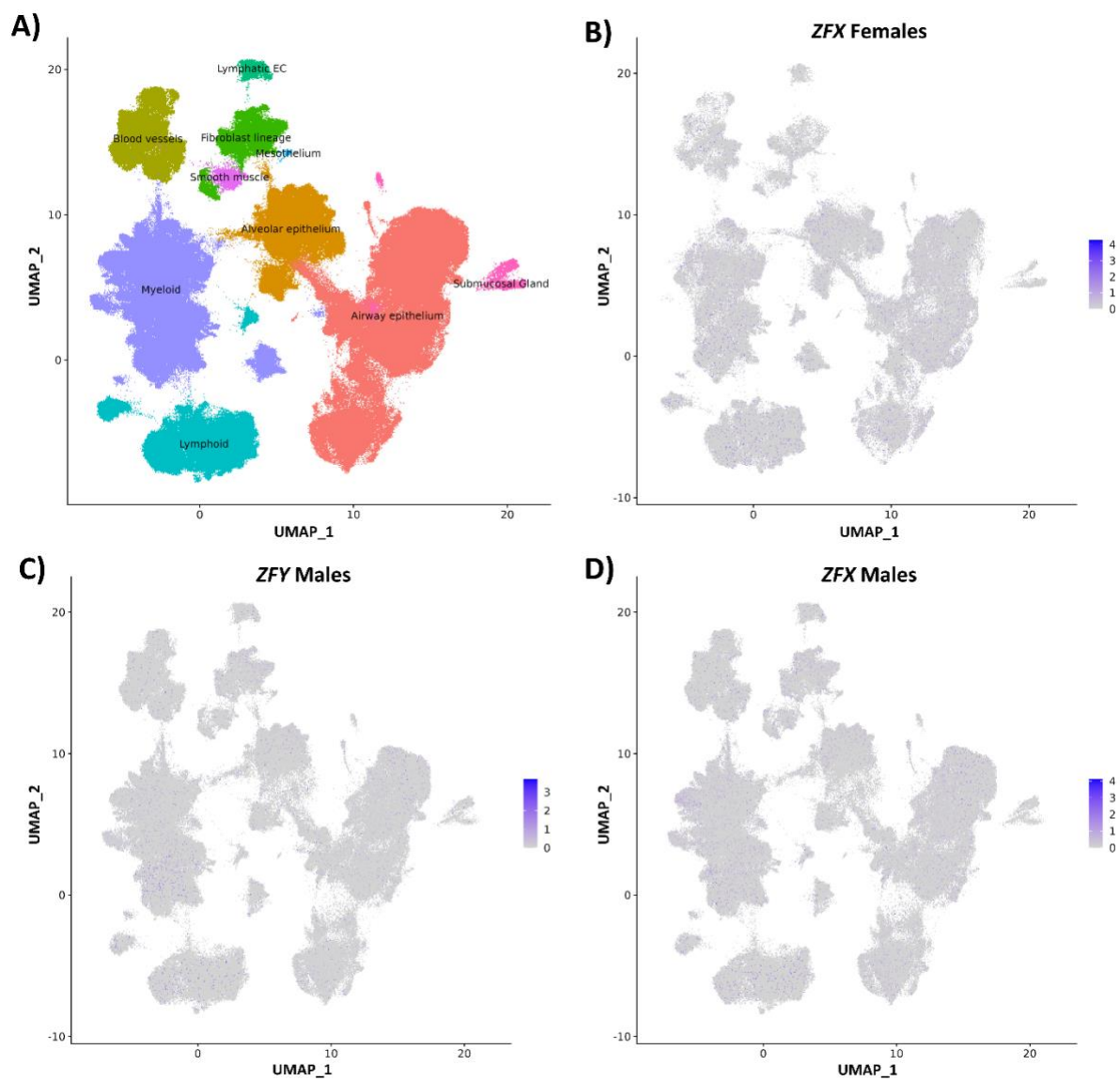
Figure S5.2: Human lung cell atlas UMAP of ZFX and ZFY gene expression. (A) Labelled UMAP of the different cell populations in the analysis. Dark purple indicates higher levels of expression of ZFX in females (B) and males (D) or ZFY in males only (C). n = 107.
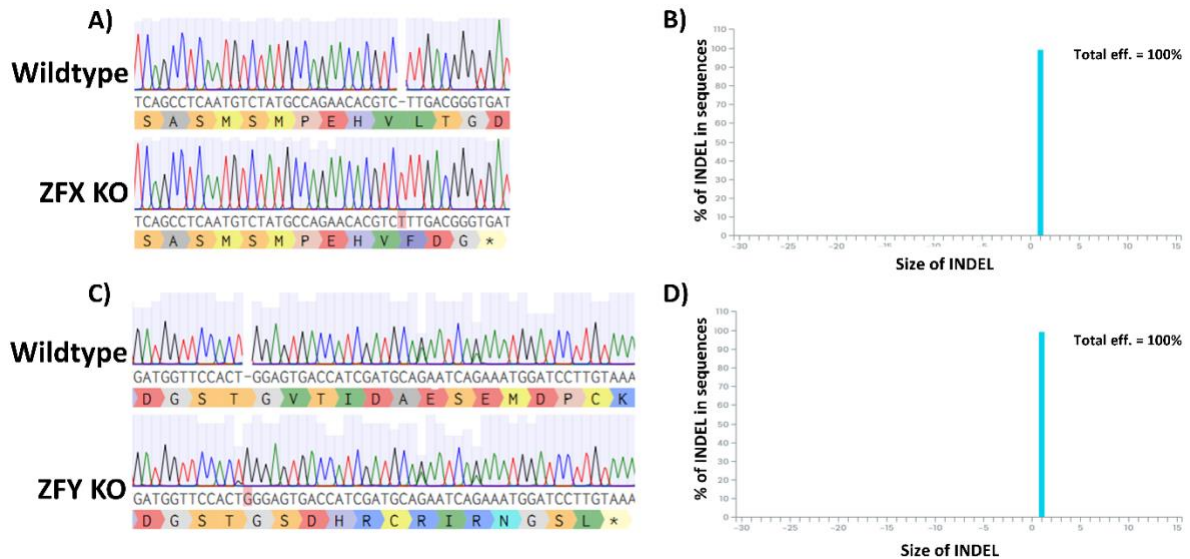
Figure S5.3: Dendrogram and indel graphic of representative of knockout cell lines. Dendrograms of ZFX (A) and ZFY (C) were generated using Benchling.com and aligned to a wildtype sequence. An inserted base is highlighted in Red, with a 'T' inserted in the representative ZFX KO and a 'G' inserted in the representative ZFY KO. The translated amino acid sequence is presented below each respective dendrogram, with an asterisk (*) indicating a stop codon. (B & D) Total eff. Indicates the percentage of the sequences in the samples that indicate the mutation shown in the dendrograms.
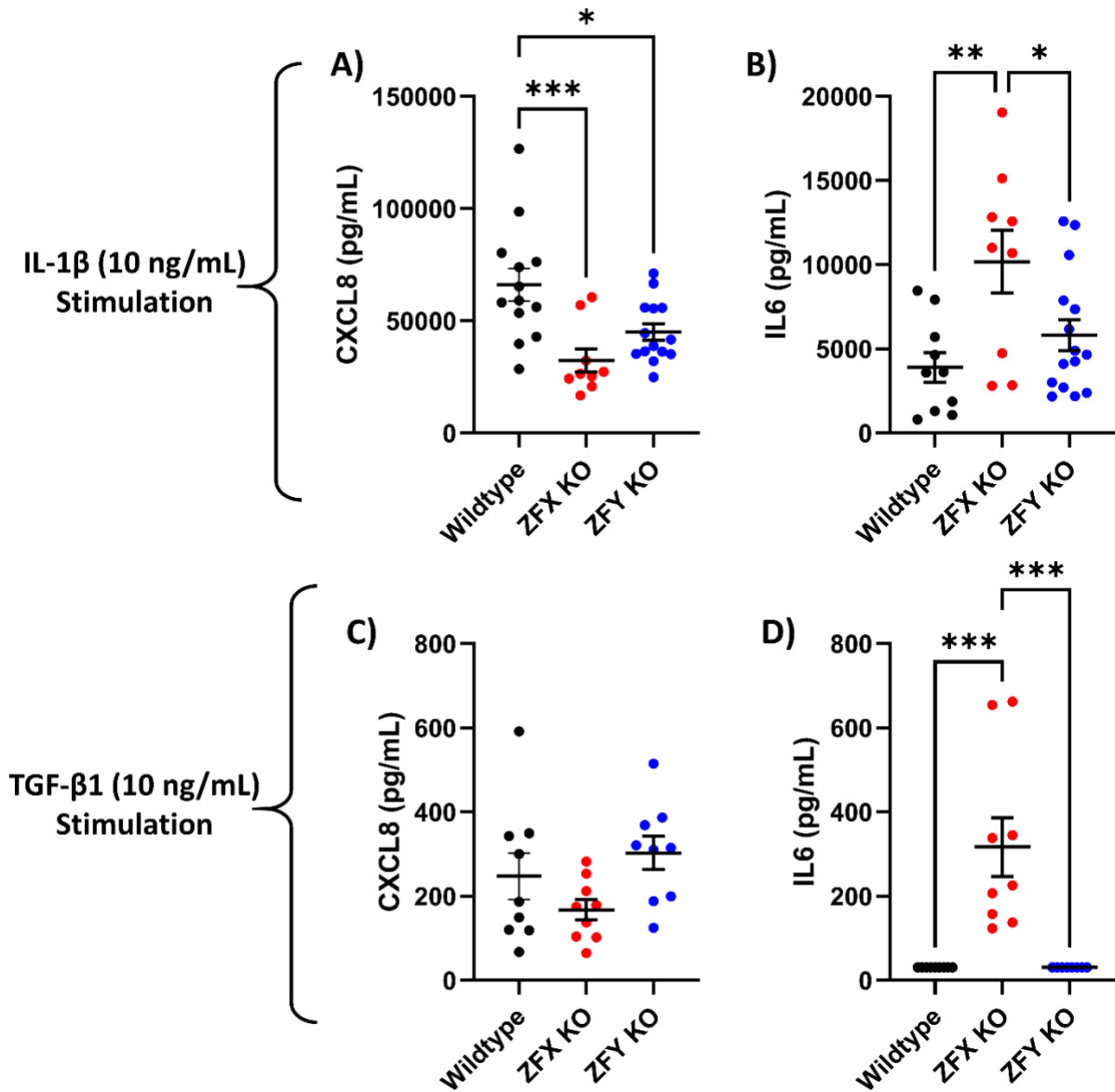
Figure S5.4: Production of CXCL8 (A & C) and IL6 (B & D) from ZFX and ZFY knockout cells when stimulated with IL-1β (A & B) and TGF-β1 (C & D). CXCL8 and IL6 were measured after 24-hour IL-1β (10 ng/mL) and TGF-β1 (10 ng/mL) stimulation in cell-free supernatant by ELISA. All data are presented as the arithmetic mean +/- SEM. One-way ANOVA statistical analysis with Tukey's correction for multiple comparisons; *p-value<0.05, **p<0.01, ***p<0.001. n= 9 – 15.

Table S5.1: Summary of g: Profiler pathway analysis using genes upregulated genes in ZFX KO and ZFY KO cell lines at baseline after 6 hours. FDR = false discovery rate

Relatively increased in ZFX KO at baseline		Relatively increased in ZFY KO at baseline	
Enriched Pathway	FDR	Enriched Pathway	FDR
Cell Migration	1.3×10^{-2}	Cytidylate kinase activity	8.3×10^{-3}
Signaling	1.6×10^{-2}	Multicellular organism development	1.5×10^{-2}
Cell Communication	1.9×10^{-2}	Developmental process	2.7×10^{-2}
Chemotaxis	2.1×10^{-2}	Response to corticotropin-releasing hormone	4.9×10^{-2}
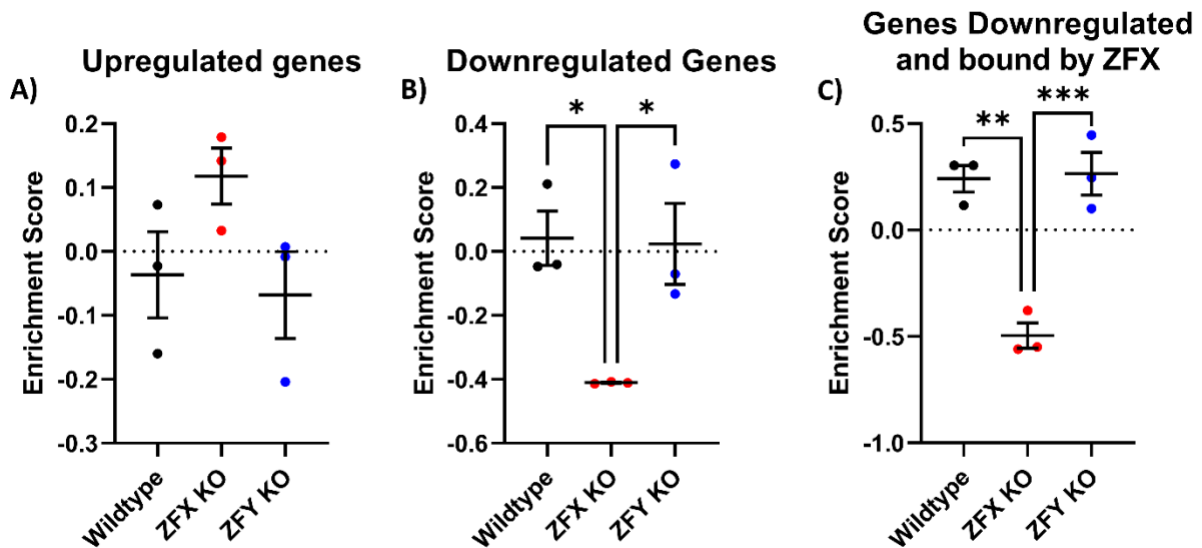
Figure S5.5: GSVA of genes up (A) or down regulated (B) in GSE145160 in knockout cell lines. GSVA enrichment score is presented on the y-axis. (C) presents genes that were downregulated in ZFX knockout cells in GSE145160 and were also shown to be binding targets of ZFX. Data are presented as the arithmetic mean +/- SEM. One-way ANOVA statistical analysis with Tukey's correction for multiple comparisons; ***$p<0.001$. n= 3.
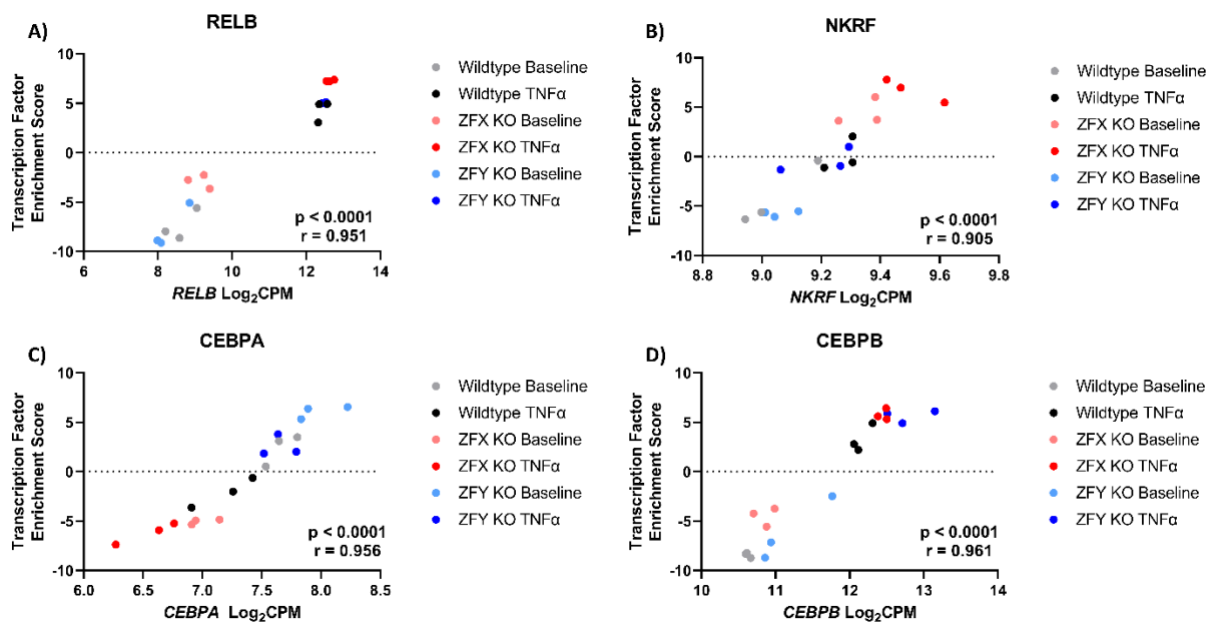
Figure S5.6: Relationship between transcription factor enrichment score (y-axis) and gene expression (x-axis). RELB (A), NKRF (B), CEBPA (C) and CEBPB (D). A two-tailed spearman correlation analysis on all samples pooled together was completed. Grey = wildtype at baseline; pink = ZFX KO at baseline; light blue = ZFY KO at baseline; black = wildtype after TNFα stimulation; red = ZFX KO after TNFα stimulation; dark blue = ZFY KO after TNFα stimulation. Cells were stimulated with TNFα (10 ng/mL) for six hours. Statistical significance was determined at p < 0.05, r = spearman coefficient; n = 3.
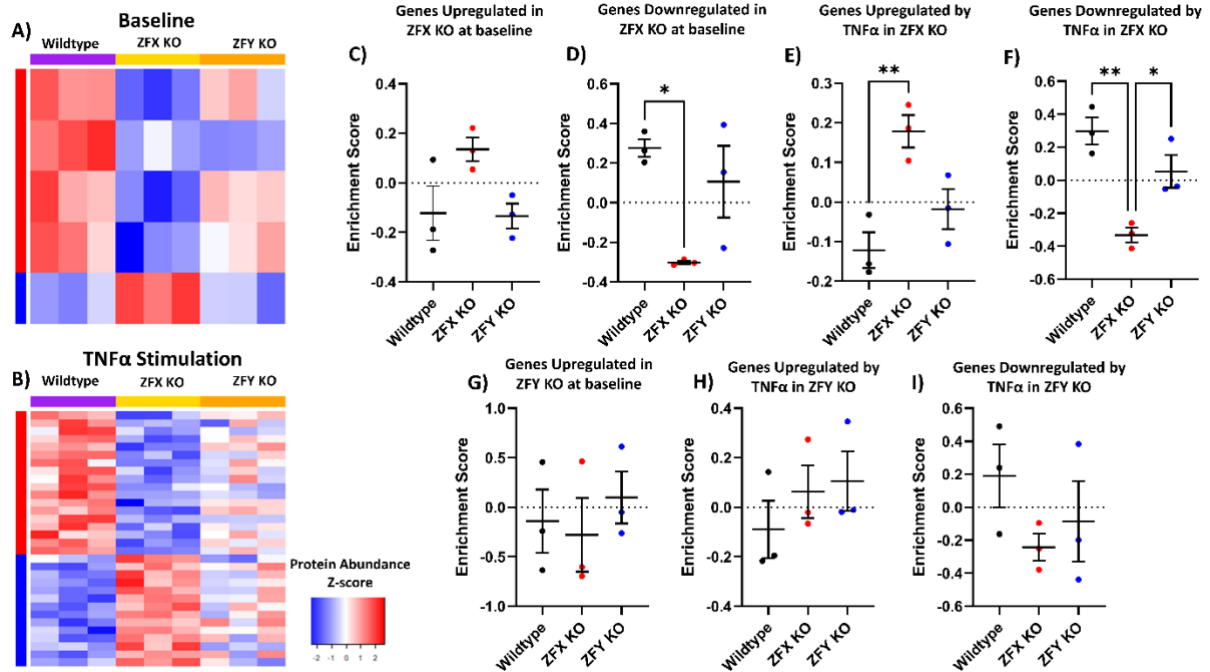
Figure S5.7: Heatmap of differentially expressed proteins between ZFX KO and wildtype cell lines and GSVA analysis of DEGs in protein dataset. A supervised heatmap of proteins differentially produced between ZFX KO cells and wildtype cells at baseline (A) and after 24-hour TNFα stimulation (B). Proteins are tracked for wildtype (purple), ZFX KO (yellow) and ZFY KO (orange). Increased genes are coloured red, and decreased genes are coloured blue. (C – I) GSVA analysis of differentially expressed genes in protein abundance dataset, data are presented as the arithmetic mean +/- SEM and analysed by one-way ANOVA with Tukey's correction for multiple comparisons. n = 3.
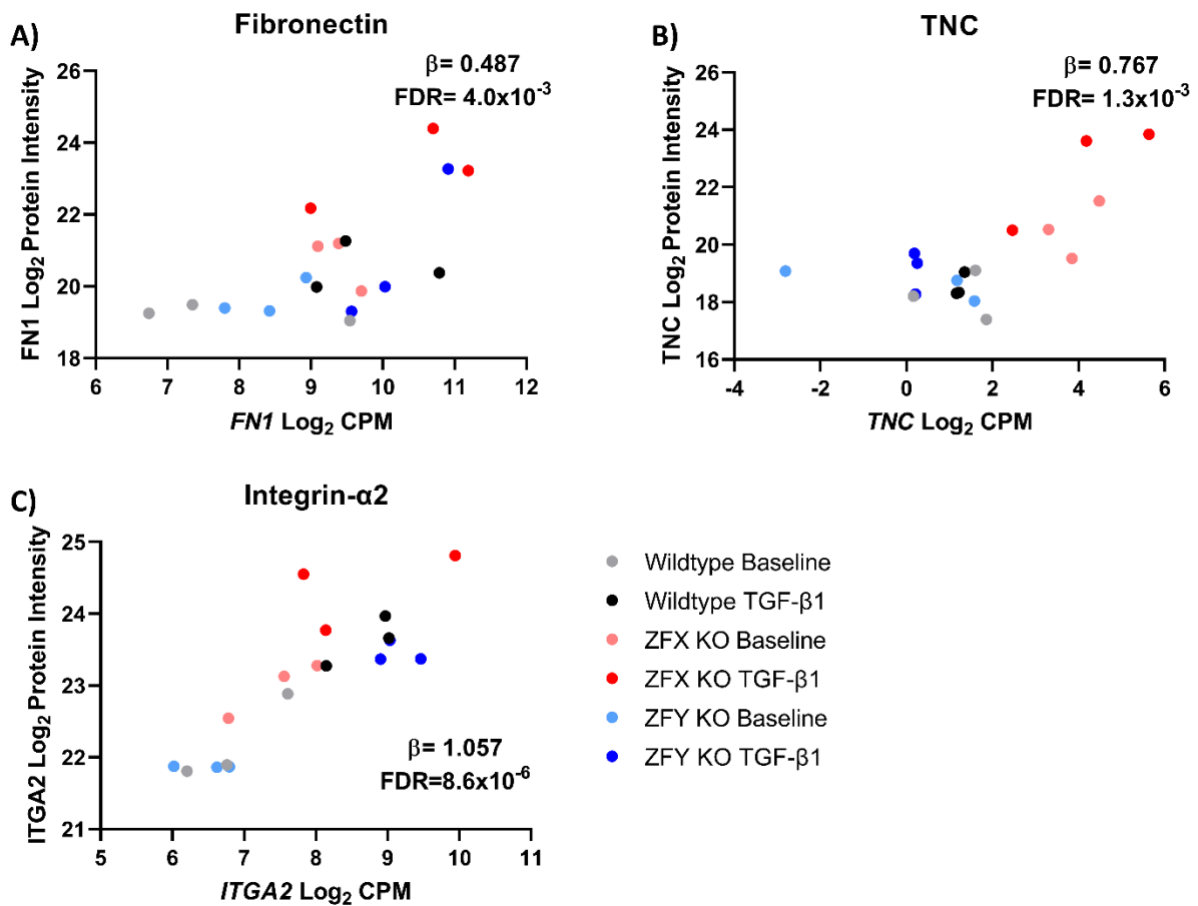
Figure S5.8: pQTL analysis of the relationship between protein abundance (y-axis) and gene expression (x-axis). (A) Fibronectin; FN1, (B) Tenascin-C; TNC and (C) Integrin-α2; ITGA2. The β-value indicates the slope of the correlation. All results returned an FDR<0.05, determined with Benjamini-Hochberg correction for multiple comparisons. n = 3.
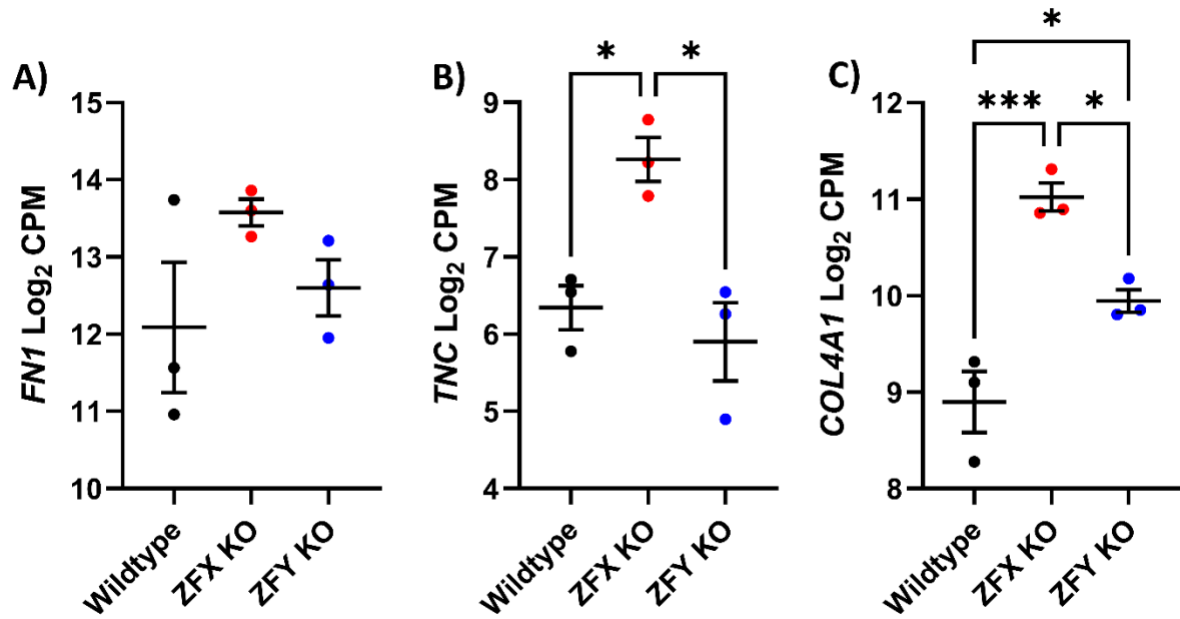
Figure S5.9: Gene expression of *fibronectin* (A), *tenascin C* (B) and *collagen 4a1* (C). Gene expression is presented as $\log_2$ counts per million (CPM) on the y-axis at baseline. Data are presented as the arithmetic mean +/- SEM and analysed by one-way ANOVA with Tukey's correction for multiple comparisons. Statistical significance is indicated by *$p<0.05$, ***$p<0.001$, n = 3.

Chapter 6 – Novel regulatory role of RPS4Y1 in inflammation and fibrotic processes

6.1 Introduction

The translation of mRNA to a polypeptide chain is essential in protein synthesis, where mRNA transcribed from DNA is converted into a functional protein. Ribosomes conduct this process. Ribosomes comprise four ribosomal RNAs (rRNAs) and multiple ribosomal proteins (RPs) [75], and these ribosomal protein subunits are differentially expressed across various tissues [76]. A dysregulation of ribosome formation can result in extra-translational functions of RPs that affect processes such as cell cycle progression, immune signalling and development [76, 77]. Studies show that RPs, rather than rRNAs, are needed for efficient ribosome assembly and mRNA translation, highlighting ribosomal proteins' critical importance.

The RPS4 ribosomal protein is an evolutionary conserved small subunit involved in mRNA binding [78]. This gene is commonly found on autosomes in vertebrate species, except in mammals where X-chromosome linked (RPS4X) and Y-chromosome linked (RPS4Y1) versions exist [78, 79]. A duplicated version of RPS4Y1 called RPS4Y2 exists but demonstrates precise expression in the testes and germ cells [80, 81], whereas RPS4Y1 and RPS4X are constitutively expressed in various tissues [81]. Further, RPS4Y1 and RPS4X are expressed in all previously studied primate lineages highlighting a vital contribution to correct development and regular cellular function [82].

RPS4X and RPS4Y1 are 263 amino acids in size but share only 92.8% amino acid homology[79], with 19 amino acid substitutions [3, 79, 83]. Lopes et al. identified eleven non-conserved amino acid substitutions [81]. These substitutions alter the chemical properties of the specific amino acid residues leading to a variation in the chemical properties between the proteins. The consequence of this difference remains unclear, although it could lead to differential functions [84]. RPS4X demonstrates complete homology with the X-chromosome linked mouse RPS4 gene [83], indicating that RPS4X structure has been conserved whilst RPS4Y1 has undergone mutations leading to a divergence from RPS4X. Therefore, as males carry both X and Y chromosomes, they express two distinct ribosomes, whilst females only express ribosomes containing RPS4X. Hence, any functional differences between RPS4X and RPS4Y1 result in sexual dimorphism at the translational level between males and females. Amino acids that are conserved between RPS4X and RPS4Y1 in humans cluster at the N-terminal domain of the protein, illustrating this region's importance for ribosomal interactions [81]. Despite amino acid substitutions, initial studies concluded that these proteins carry essential functions in translationally active ribosomes [79, 84]. These findings demonstrate ribosomes containing either RPS4X or RPS4Y1 but do not explore whether RPS4Y1-

containing ribosomes carry a unique function in males. Importantly, RPS4Y1-containing ribosomes are 10-15% as abundant as RPS4X-containing ribosomes [83]. Amino acid substitutions between the protein isoforms are predicted to make RPS4Y1-containing ribosomes more basic [83]. Different pH levels have been shown to alter ribosomal function[85, 86], potentially causing preferential translation of specific gene transcripts. As such, RPS4Y1 vs RPS4X-containing ribosomes may have distinct functions. This notion may contribute to sex-biased gene expression [2], subsequently driving sex differences in disease susceptibility.

Many human diseases demonstrate sexual dimorphism in development, progression and prognosis. Asthma is a chronic inflammatory respiratory disease that culminates in variable airflow in patients. It affects nearly 300 million people worldwide [31], yet the exact cause of asthma remains unknown. Importantly, asthma demonstrates a distinct sexually dimorphic pattern of susceptibility and severity [32]. In childhood, males have both increased incidence and worse asthma outcomes [87]. However, females show greater susceptibility to asthma development in adulthood and present with a higher hospitalisation rate [30]. Several factors have been proposed to contribute to the complex interaction between biological sex and asthma, such as the sex hormones, physiological differences and the sex chromosome complement [28].

The gene expression and dosage imbalance due to sex chromosome complement between males and females has garnered greater attention as a critical biological factor contributing to disease processes [40, 88]. In particular, RPS4X has been implicated with worse prognosis in various cancer types [89-91], which is likely linked to RPS4X regulation of cell proliferation[89]. RPS4Y1 has been less widely studied but has been identified to be dysregulated in asthma [92]. Further, RPS4Y1 regulates the response of patients to oral corticosteroids [93] (an essential treatment for asthma patients), modulation of the inflammatory response [94] and fibrotic processes [95]. Notably, these processes all form critical, hallmark features of asthma, highlighting a potential contribution of RPS4Y1 to the pathology of asthma.

Studies investigating the relationship between RPS4Y1 and disease processes do not examine the molecular mechanisms regulated by RPS4Y1. This study aims to explore and characterise the regulatory role of RPS4Y1 for inflammatory and fibrotic processes relevant to asthma. Through a multi-omics approach, we highlight RPS4Y1 functions via distinct and variable mechanisms to regulate protein production. We demonstrate RPS4Y1 modulates the production of inflammatory factors and critical extracellular matrix proteins, affecting cell adhesion and migration. In particular, we identify the potential for RSP4Y1 to preferentially increase the translation of IL6 and tenascin-C mRNA. These data generate an RPS4Y1-

specific gene signature correlating with asthma severity in male patients determined via lung function. Importantly, our study sheds light on the novel regulatory function of RPS4Y1. Females do not express RPS4Y1, so its functions may contribute to sex differences in asthma susceptibility and progression.

6.2 Methodology

6.2.1 Analysis of *RPS4X* and *RPS4Y1* expression in healthy and asthmatic patients

The publicly available single-cell RNA-seq 'Human Lung Cell Atlas'[15] and 'Lung Cell Atlas' [16] were accessed to analyse the gene expression of *RPS4X* and *RPS4Y1* in different epithelial cell types in non-asthmatic and asthmatic patients[15]. The Indurian and OLiVIA RNA-seq datasets were also used to assess *RPS4X* and *RPS4Y1* expression in males and females. Patient demographics for all studies are located in section 3.16.

6.2.2 Investigation of RPS4X and RPS4Y1 essentiality for cell survival using DepMap portal

The publicly available Cancer Dependency Map (DepMAP) portal (https://depmap.org/portal/interactive/) was used to investigate the 'essentiality' of the *RPS4X* and *RPS4Y1* genes for cell survival. All lung adenocarcinoma cell lines were selected with the 'essentiality' score as determined in the *CRISPR Chronos* dataset. A score less than negative indicates that the gene is essential for survival, while a score greater than zero indicates that the gene contributes to proliferation.

6.2.3 Generation of CRISPR Cas9 genome deletion cell lines

Three unique RPS4Y1 knockout cell lines were established in A549 cells, as described in section 3.2. A schematic representation of the overall study design and analyses is included in the supplement as Supplementary Figure S6.1.

6.2.4 Western blot

RPS4Y1 knockouts were confirmed by western blot analysis. Anti-human RPS4Y1 (#17296-1-AP, ProteinTec) was used to detect and visualise protein bands. Anti-human GAPDH was used as a loading control (#MAB374, Merck-Millipore). Detailed methodology is located in section 3.19.

6.2.5 Cell culture and treatments

All generated cell lines were maintained in DMEM growth medium. A detailed description of cell maintenance techniques and cell treatments are located in section 3.1.

6.2.6 Measurement of CXCL8 and IL6 protein secretion

Wildtype and RPS4Y1 KO cell-free supernatants were collected 24 hours post-stimulation with TNFα (10 ng/mL), IL-1β (10 ng/mL) and TGF-β1 (10 ng/mL). The concentration of CXCL8 and IL6 was quantified by ELISA and was completed as described in section 3.3.

6.2.7 RNA-sequencing

Whole-cell RNA extracts from wildtype and RPS4Y1 KO cells were collected at 6 hours post-TNFα (10 ng/mL) and 48 hours post-TGF-β1 (10 ng/mL) stimulation and processed as described in section 3.10 and section 3.11.

6.2.8 Differential gene expression analysis

Differential gene expression analysis was completed using the *Dseq2* package in R. RPS4Y1 KO cell gene expression was compared to wildtype cell gene expression at baseline, after 6-hour TNFα (10 ng/ml) and 48-hour TGF-β1 (10 ng/ml) stimulation. A detailed description of differential gene expression (DGE) analysis is located in section 3.12.

6.2.9 LC-MS/MS proteomics analysis

Proteomics analysis was completed on wildtype and RPS4Y1 KO cell lines as described in section 3.17. Protein lysate samples were collected at 24 hours post-TNFα (10 ng/mL) and 72 hours post-TGF-β1 (10 ng/mL) stimulation.

6.2.10 Transcription factor enrichment analysis

Transcription factor analysis was completed for wildtype and RPS4Y1 KO samples at baseline and after 6-hour of TNFα (10 ng/ml) stimulation as described in section 3.13.

6.2.11 Gene set variation analysis (GSVA)

GSVA was completed as described in section 3.14. This analysis was completed using the protein abundance dataset generated by LC-MS/MS. This analysis was conducted to compare the cell lines' transcriptome and proteome by tracking the production of DEGs to their corresponding proteins in the proteomics analysis dataset.

GSVA was also completed on the bronchial biopsies collected from asthmatic and healthy control patients from the Indurian dataset to analyse the association between knockout cell line gene expression and clinical phenotypes.

6.2.12 Analysis of biological pathways enriched in knockout cell lines

Pathway analysis was completed using the g: Profiler online tool on differentially expressed genes between RPS4Y1 KOs and wildtype cells, as described in section 3.15.

6.2.13 Wound healing assay

The wound healing assay was completed on wildtype and RPS4Y1 KO cells over 72 hours, as described in section 3.4.

6.2.14 Cell proliferation

The cellular proliferation rate of wildtype and RPS4Y1 KO cells was analysed by manual cell counting after 96-hour of incubation in growth medium, as described in section 3.9.

6.2.15 Cell adhesion to fibronectin

The analysis of fibronectin-mediated cellular adhesion of wildtype and RPS4Y1 KO cells after one-hour incubation in DMEM supplemented with 1% (v/v) FBS was completed, as described in section 3.5.

6.2.16 Measurement of ECM protein

Assessment of changes in the deposition of the extracellular matrix proteins fibronectin (FN1), tenascin-C (TNC) and collagen 4α1 (COL4A1) from wildtype and RPS4Y1 KO cells after 72-hour incubation was completed as described in section 3.8.

6.2.17 Cigarette smoke extract (CSE) generation and analysis cell death response

Marlboro Red standard cigarettes (Philip Morris) were used to prepare CSE, as described in section 3.6.

6.2.18 Cell viability assay

The cellular viability of wildtype and RPS4Y1 KO cells after 24 and 48-hour CSE exposure was assessed using the MTT assay as described in section 3.7.

6.2.19 Protein quantitative trait loci (pQTL) validation

Gene expression correlation with protein abundance levels was analysed using protein quantitative trait loci for wildtype and RPS4Y1 KO cells as described in section 3.18.

6.3 Results

6.3.1 *RPS4Y1* expression is differentially regulated in asthma

To explore the expression profile of *RPS4Y1* and *RPS4X,* we investigated their expression at a single-cell level and in the lung of healthy and asthmatic patients. The heatmaps present the expression of *RPS4X* (Figure 6.1A) and *RPS4Y1* (Figure 6.1B) across multiple cell types and indicate how their expression changes with asthma diagnosis. Both heatmaps demonstrate that *RPS4X* and *RPS4Y1* gene expression varies depending on the cell type but also changes with disease status. Of note, *RPS4X,* although different between cell types and most expressed in 'basal activated' cells, indicates limited changes with asthma status. *RPS4Y1* show negligible expression of 'basal activated' cells in healthy patients but is significantly upregulated in these cells with asthma. As such, regulating *RPS4Y1* expression demonstrates significant responsiveness in asthma, particularly in basal epithelial cells. The UMAP analysis supports functions for these genes in basal cells. *RPS4X* is more widely expressed at higher levels than RPS4Y1, although both genes are expressed highly from the alveolar and airway epithelium in males and females (Figure 6.1G – 6.1J).

Following the observations at a single-cell level, we expanded our analysis to the respiratory tract. We explored whether the expression of *RPS4X* and *RPS4Y1* in males equivocated the expression of *RPS4X* from both X-chromosomes in females. In nasal brushings, *RPS4X* expression from females was greater than the combined expression in males (Figure 6.1C). However, in the bronchus, the expression levels between males and females were equal (Figure 6.1D). Importantly, in both the nasal and bronchial samples, RPS4X expression comprised approximately 80% of the total expression in males. Finally, we confirm in bronchial biopsies that *RPS4Y1* is downregulated in males with asthma (Figure 6.1E), but RPS4X demonstrates consistent expression between healthy and asthmatic patients from both sexes (Figure 6.1F). These data highlight that *RPS4Y1* is dysregulated in asthma. Therefore, it is essential to characterise its contribution to the hallmark features of asthma.
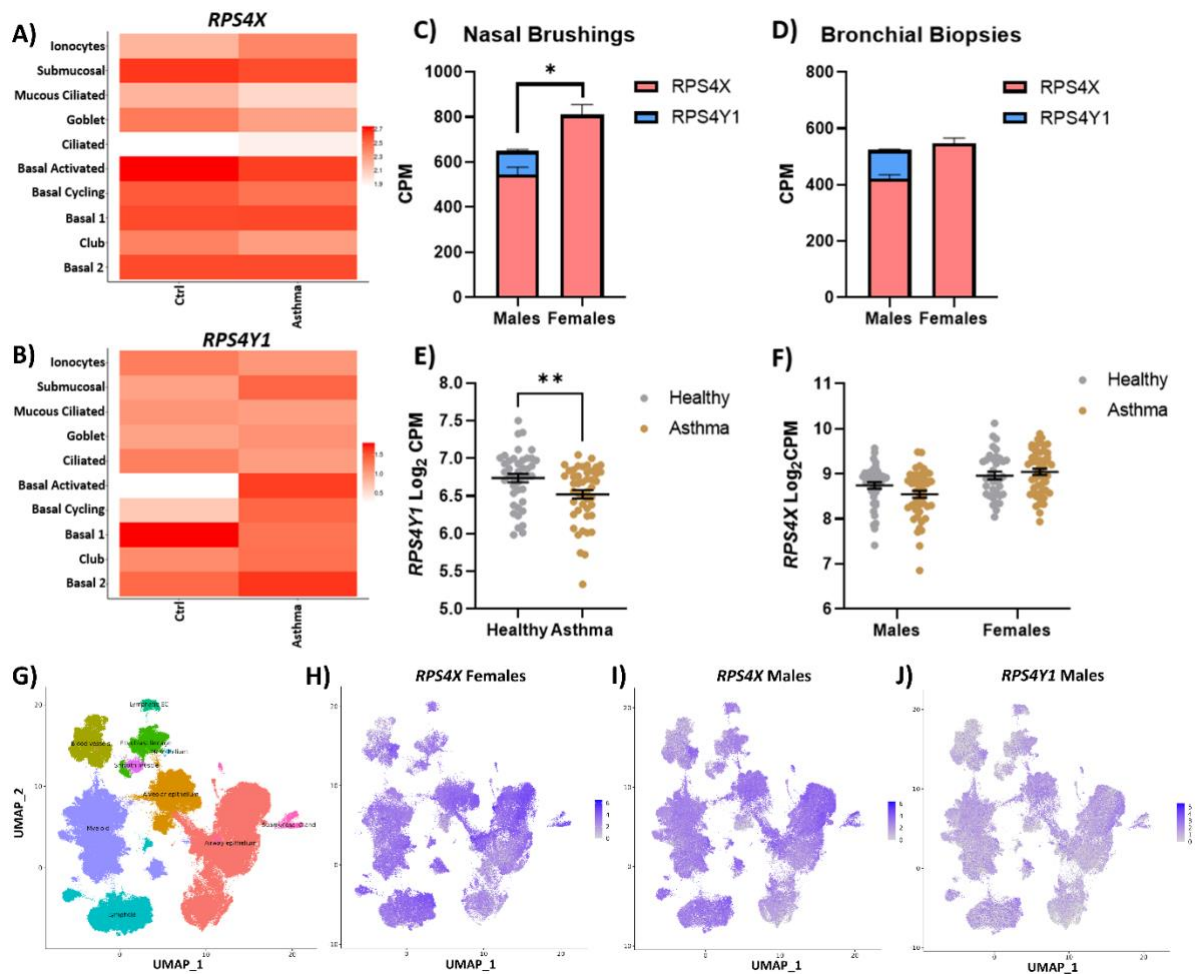
Figure 6.1: Gene expression of *RPS4Y1* and *RPS4X* at a single cell level, bronchial biopsies and nasal brushings. Heatmaps of single-cell sequencing expression of *RPS4X* (A) and *RPS4Y1* (B) across multiple cell types in healthy control (ctrl); n = 4 female/ 7 males, and asthma; n = 2 female/ 7 male patients. Dark red represents higher expression, and white equals no expression. (C) nasal brushing; n = 101 females/ 89 males and (D) bronchial biopsy; n = 85 females/ 88 males of RPS4X (pink) and RPS4Y1 (blue) expression as counts per million (CPM). Data is analysed by fitting a mixed effects model. (E) *RPS4Y1* expression as $\log_2$ CPM in males healthy and asthmatic bronchial biopsies. (F) *RPS4X* expression as $\log_2$ CPM in bronchial biopsies from healthy and asthmatic patients of both sexes. Grey represents healthy patients, and gold represents asthma patients. (C – F) data are presented as the arithmetic mean +/- SEM. (E & F) is analysed by unpaired parametric t-test, statistical significance is indicated by *$p<0.05$ and **$p<0.01$. UMAP of the identified cell-types of single-cell RNA-sequencing human lung cell atlas. (G) UMAP of defined cell locations, *RPS4X* expression in females (H) and males (I), (J) *RPS4Y1* expression in males. Purple indicates higher expression, and white indicates lower expression.

6.3.2 RPS4Y1 regulates inflammatory cytokines and cellular attachment

The UMAPs (Figure 6.1G – 6.1J) highlight that *RPS4Y1* is expressed from alveolar and airway epithelial cells and is dysregulated in asthma (Figure 6.1B). As such, three unique RPS4Y1 knockout (KO) cell lines were generated in an alveolar basal epithelial cell line (A549) using CRISPR-Cas9. These KO cell lines were characterised by changes in the inflammatory response, fibrotic processes and the regulation of cell death. Successful RPS4Y1 KO was confirmed by genome sequencing (Figure 6.2A), *RPS4Y1* gene expression (Figure 6.2B) and western blot (Figure 6.2D). *RPS4X* gene expression significantly increases in RPS4Y1 KO cells (Figure 6.2C). This highlights a specific regulatory feedback mechanism between RPS4Y1 and RPS4X, where RPS4X production increases to compensate for the loss of RPS4Y1. It was not possible to generate RPS4X KO cell lines. Analysis of the online DepMap dataset revealed that RPS4X is essential for cell survival, as indicated by an 'essentiality' score less than negative one (Figure 6.2E).

The immunoregulatory contribution of RPS4Y1 was assessed by stimulating cells with TNFα for 24 hours. In RPS4Y1 KOs, CXCL8 production significantly decreases (Figure 6.2F), whilst IL6 output is significantly increased (Figure 6.2G). The cells were also stimulated with IL-1β and TGF-β1 (two other pathologically relevant inflammatory cytokines), where the same pattern of regulation for CXCL8 and IL6 was observed (Supplementary Figure S6.2). This highlights an essential role for RPS4Y1 in regulating these cytokines, which have different regulatory pathways.

RPS4Y1 KO cells demonstrated a significantly faster wound closure rate than wildtype cells (Figure 6.2H). We investigated cell attachment and cell proliferation to explore possible factors contributing to this observation. RPS4Y1 KOs reported increased cellular adhesion to fibronectin (Figure 6.2I), whilst no difference in proliferation to wildtype cells is noted (Figure 6.2J). Combined, these results indicate that increased adhesion promotes increased migration of RPS4Y1 KO cells leading to a faster rate of wound closure. No change in cell survival to cigarette smoke extract (CSE) was observed (Figure 6.2K).

These data demonstrate that RPS4Y1 has a complex contribution to regulating asthma-relevant disease processes of inflammation, cell migration and adhesion, and changes to the extracellular matrix.
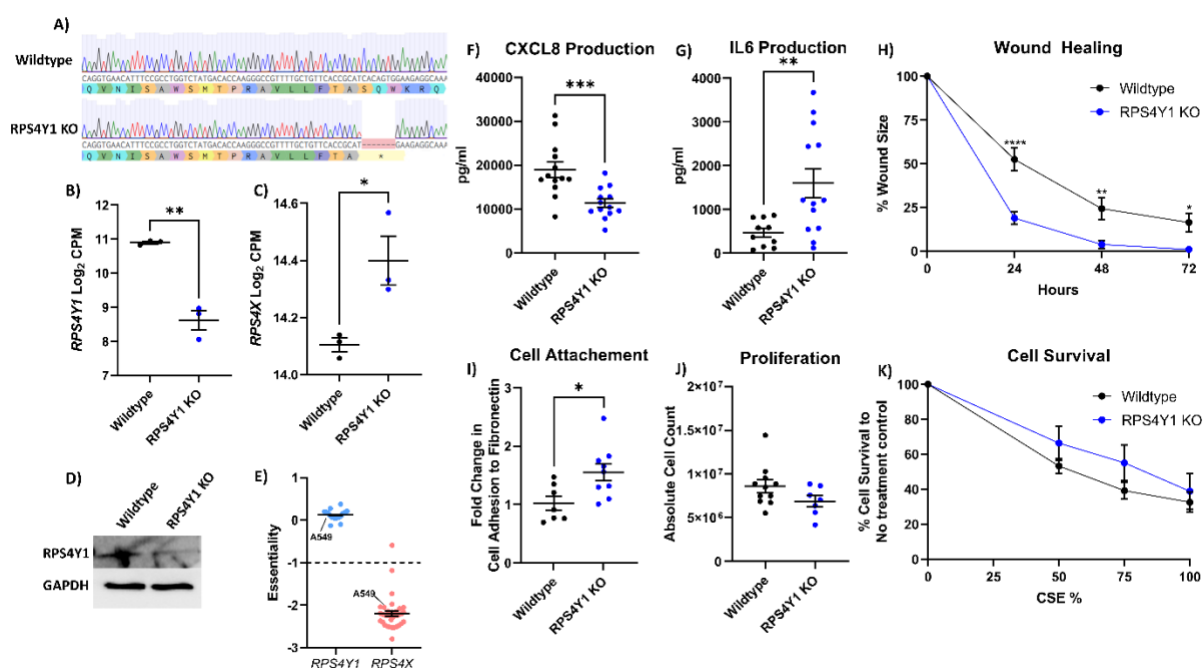
Figure 6.2: Characterisation of RPS4Y1 knockout cells (blue dots) compared to wildtype cells (black dots). (A) Representative chromatogram of wildtype and RPS4Y1 genome sequencing with codon translation to amino acid sequence underneath. Asterisk (*) indicates a stop codon. Log_2 counts per million (CPM) expression of *RPS4Y1* (B) and *RPS4X* (C) was measured by RNA-sequencing. (D) Representative western blot image confirming RPS4Y1 KO. (E) Essentiality score for *RPS4Y1* and *RPS4X* in multiple lung cell lines with A549 cells identified; score < -1 indicates the gene is essential for survival. CXCL8 (F) and IL6 (G) production was measured in cell-free supernatant after 24-hour stimulation with TNFα (10 ng/mL) by ELISA. (H) The size of a scratch wound was measured every 24 hours for three days with incubation in growth medium. Wound closure was measured as a percentage compared to the initial wound size at zero hours. (I) Cell attachment to fibronectin was measured after 1 hour of incubation. (J) Proliferation was measured after 96 hours of incubation in growth medium by manual cell counting. (K) Percentage cell survival compared to no treatment control (y-axis) was determined by an MTT assay after 24-hour cigarette smoke extract (CSE) exposure. All data are presented as the arithmetic mean +/- standard error of the mean. (H & K) analysed by two-way ANOVA with Sidak correction for multiple comparisons testing. All other analyses were analysed by unpaired parametric t–tests. Statistical significance is represented by *$p<0.05$, **$p<0.01$, ***$p<0.001$ and ****$p<0.0001$.

6.3.3 RPS4Y1 knockout significantly alters gene expression

Now that a contribution of RPS4Y1 was established at the phenotypical level, we conducted RNA-sequencing to determine whether knocking out RPS4Y1 significantly alters gene expression. Sequencing was conducted after 6 hours of incubation with or without TNFα stimulation. At baseline, RPS4Y1 KO cells demonstrated 125 significantly increased genes (FDR < 0.05 and log_2FC > 1.0) and 139 significantly downregulated genes (FDR < 0.05 and log_2FC < -1.0). An interaction analysis was completed, comparing genes differentially regulated between wildtype and KO cell lines after TNFα stimulation (Figure 6.3A). The interaction analysis revealed no significantly upregulated genes, whilst *SHISA9* and *SPDEF* were the only genes downregulated (Figure 6.3B). This means these two genes are relatively less expressed in RPS4Y1 KOs than wildtype cells after TNFα stimulation. *SHISA9* expression is significantly increased in wildtype cells upon TNFα stimulation, but this is lost in RPS4Y1 KO cells (Figure 6.3F). *SPEDEF* is an integral immunoregulatory and goblet cell differentiation factor in airway diseases[96]. This gene demonstrates reduced expression in RPS4Y1 KOs at baseline and a greater decrease in expression in knockout cells after TNFα stimulation (Figure 1G). The heatmap in Figure 6.3C highlights the distinct transcriptional profiles between wildtype and knockout cells at baseline.

Of particular interest is the expression of CXCL8 and IL6 after TNFα stimulation. In RPS4Y1 KOs, the expression of CXCL8 is downregulated (Figure 6.3D). This correlates with the observed suppression of CXCL8 protein production in Figure 6.2F. Conversely, IL6 expression is consistent between wildtype and KO cells (Figure 6.3E), in contrast to the increased production of IL6 protein reported in Figure 6.2G. RPS4Y1 is a ribosomal protein that contributed to mRNA translation to protein. Our results indicate that in the absence of RPS4Y1, translation of IL6 mRNA into protein might preferentially increase in response to proinflammatory stimuli.
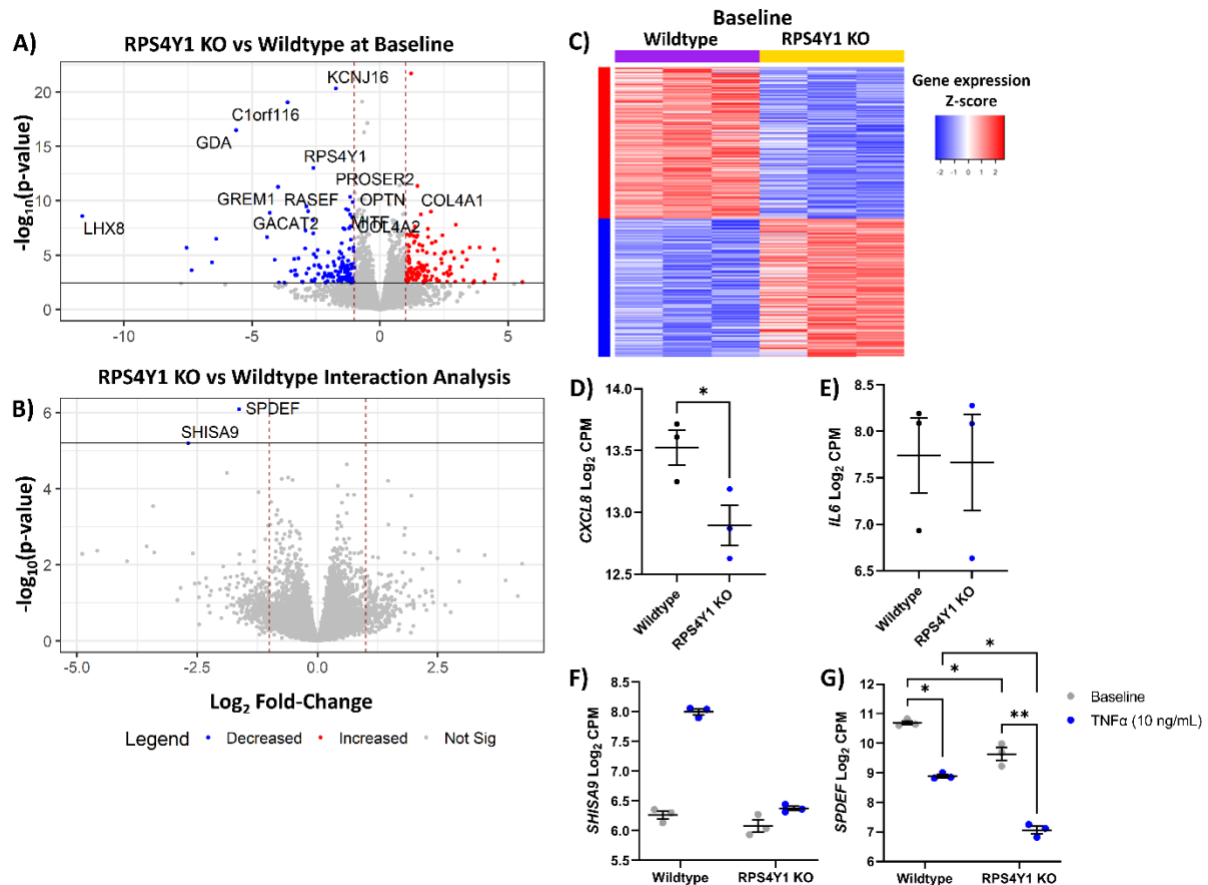
Figure 6.3: Differentially gene expression analysis of RPS4Y1 knockout vs. wildtype cell lines. Volcano plots of $-\log_{10}$(p – value) against $\log_2$fold-change in differentially expressed genes (DEGs) at baseline (A) and when KO and wildtype cell lines are stimulated with TNFα (10 ng/mL) for 6 hours minus baseline DEGs (B). The horizontal black line represented a false discovery rate < 0.05; dotted red vertical lines indicate $\log_2$fold-change > |1.0|. Red dots indicate significantly upregulated genes, and blue dots represent significantly downregulated genes in RPS4Y1 KOs. (C) Semi-supervised heatmap of DEGs at baseline between wildtype (purple) and RPS4Y1 KO (yellow) cell lines. Red indicates increased expression, and blue indicates reduced expression. $\log_2$ counts per million (CPM) expression of *CXCL8* (D), *IL6* (E), *SHISA9* (F) and *SPDEF* (G) in wildtype and RPS4Y1 KO cells. Grey dots represent baseline levels; black and dark blue dots represent TNFα (10 ng/mL) stimulated samples. (A – E) Statistical significance was determined through differential gene expression analysis with multiple comparisons correction completed using the Benjamini-Hochberg method. (F & G) Two-way ANOVA completed statistical analysis with Tukey's correction for multiple comparison testing. (D – E) Data are presented as the arithmetic mean +/- SEM. Statistical significance is represented by *p<0.05, **p<0.01. n = 3 for all analyses.

6.3.4 RPS4Y1 regulates specific inflammatory and migration-related pathways

To explore the pathways contributing to CXCL8 and cell adhesion, we analysed RPS4Y1 KOs at a protein level using LC-MS/MS. This analysis revealed that with 24-hour TNFα stimulation, three proteins were detected to be reduced in KO cells, with 20 individual proteins increased when RPS4Y1 is knocked out (FDR < 0.05, $\log_2$FC > 1.0, Figure 6.4A). NF-κB-repressing factor (NKRF) and syntenin-1 (SDCBP) were essential proteins raised in RPS4Y1 KOs (Figure 6.4B & 6.4C). Higher levels of NKRF are linked to reduced CXCL8 production [57]. Significantly, NKRF gene expression does not correlate with protein production (Figure 6.4D). This indicates that the absence of RPS4Y1 enables more efficient translation of NKRF mRNA into protein, which may contribute to the suppression of CXCL8.

To further investigate the contribution of RPS4Y1 to the regulation of gene expression, we completed a transcription factor enrichment analysis (Figure 6.4E). This identified multiple differentially activated transcription factor-related pathways. Notably, FOXA1 (Figure 6.4F) and STAT3 (Figure 6.4G) were negatively enriched in RPS4Y1 KO cells. As such, genes regulated by these genes demonstrate an overall reduction compared to wildtype cells. Studies have shown that FOXA1 and STAT3 activation significantly increase CXCL8 secretion, and IL8 mRNA levels are reduced when both factors are knocked down. Therefore, reduced activation of the FOXA1 and STAT3 pathways may be driving the suppression of CXCL8 production in combination with increased levels of NKRF.

SDCBP is increased at a protein level (Figure 6.4C), with overexpression of this protein linked to the promotion of cell migration and invasion [97]. Pathway analysis using g: Profiler revealed that cell migration and motility are positively enriched in RPS4Y1 KOs, where integrins mediate increased adhesion in RPS4Y1 KOs (Table 6.1). In support, we observe increased expression of integrin subunit α4 (*ITGA4*) and decreased expression of integrin subunit β8 (*ITGB8*) in RPS4Y1 KOs (Figure 6.4H & 6.4I). The relative expression reported for these integrin molecules has been linked with increased adhesion to fibronectin and more significant cell migration.
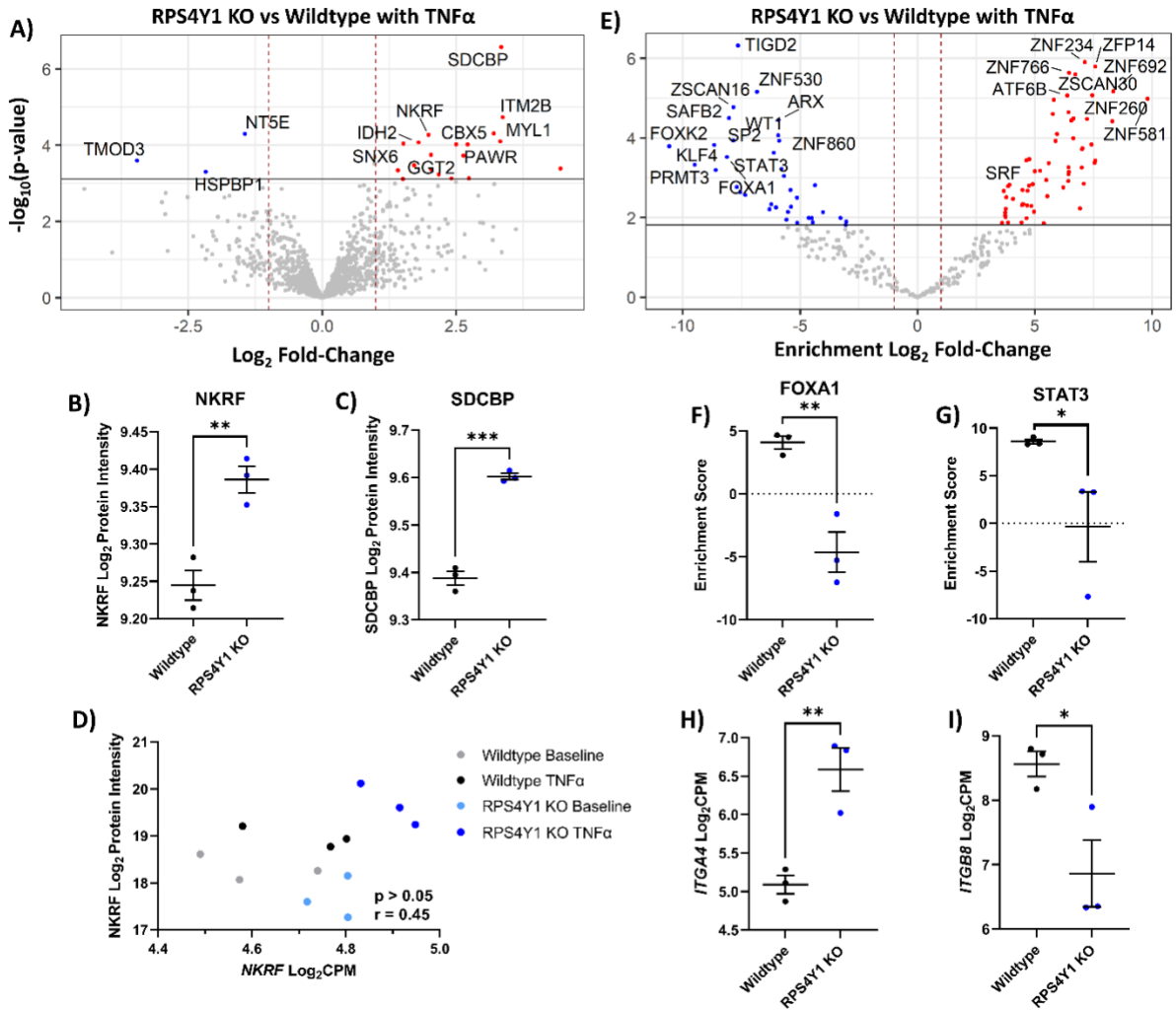
Figure 6.4: Differential protein expression and transcription factor enrichment analysis after TNFα stimulation of wildtype and RPS4Y1 KO cell lines. (A) Volcano plot of $\log_{10}$ (p – value) against $\log_2$ fold protein abundance change after 24-hour TNFα-stimulation (10 ng/mL). $\log_2$ protein intensity determined by LC-MS/MS for NKRF (B) and SDCBP (C). (D) pQTL linear regression analysis of NKRF $\log_2$ protein intensity against *NKRF* $\log_2$ counts per million (CPM) gene expression in wildtype and RPS4Y1 KO cells. (E) Volcano plot of $-\log_{10}$ (p – value) against $\log_2$ fold-change in transcription factor enrichment score. For both volcano plots (A & E), horizontal black lines represented a false discovery rate < 0.05, and dotted red vertical lines indicate $\log_2$ fold-change > |1.0|. Red dots indicate increased production (A) or positive enrichment (E), and blue dots indicate reduced production (A) or negative enrichment (E) in RPS4Y1 KOs. (F & G) transcription factor enrichment score for FOXA1 and STAT3. $\log_2$ CPM for *ITGA4* (H) and *ITGB8* (I) in wildtype and RPS4Y1 KOs. (B, C, F – I) Data presented as the arithmetic mean +/- SEM and analysed by unpaired parametric t – test. (D) Analysed by a spearman correlation with 'r' representing the correlation coefficient. Statistical significance indicated by *p<0.05, **p<0.01 and ***p<0.001. n = 3 for all analyses.

Table 6.1: Summary of g: Profiler pathway analysis using significantly up and downregulated genes in RPS4Y1 KO cells compared to wildtype cells at baseline.

Positively Enriched in RPS4Y1 KO at Baseline		Negatively Enriched in RPS4Y1 KO at Baseline	
Pathway	**FDR**	**Pathway**	**FDR**
Regulation of cell migration	4.33×10^{7}	Actin binding	1.47×10^{-2}
Regulation of cell motility	1.05×10^{7}	Guanine metabolic process	3.56×10^{-2}
Regulation of signalling	2.21×10^{-3}	Myosin complex	4.71×10^{-2}
Cell adhesion mediated by integrins	2.36×10^{-3}		

FDR = false discovery rate

6.3.5 RPS4Y1 mediates the expression and translation of specific extracellular matrix proteins

The extracellular matrix (ECM) and its integral proteins are important factors that mediate cell proliferation, migration and airway remodelling in asthma. Changes to these proteins are prominent in asthma and contribute to a worse disease prognosis. As such, we investigated how asthma-associated ECM proteins fibronectin (FN1), tenascin-C (TNC) and collagen 4α1 (COL4A1) are altered at a transcriptional and protein level in the absence of RPS4Y1. Differential gene expression analysis after 48-hour incubation after quiescence untreated, revealed 1315 differentially expressed genes between RPS4Y1 KOs and wildtype cells at baseline (Figure 6.5A). To explore the correlation between changes in gene expression and protein production, GSVA was used to analyse how differentially expressed genes are changed in the LC–MS/MS proteomics dataset. This analysis indicated that genes upregulated or downregulated in RPS4Y1 KOs at baselines were similarly altered at a protein level (Figure 6.5B & 6.5C).

Amongst the differentially expressed genes identified in Figure 5A, fibronectin and collagen 4α1 gene expression is significantly increased (Figure 6.5D & 6.5F). This increase in mRNA expression is reflected at a protein level measured by ECM ELISA. Fibronectin indicates a trend towards increased production from RPS4Y1 KO cells (Figure 6.5E), whilst collagen 4α1 production from KO cells is approximately double what wildtype cells produce (Figure 6.5G). Conversely, tenascin-C indicates similar expression between wildtype and RPS4Y1 KO cell lines (Figure 6.5H) yet demonstrates significantly increased production at a protein level when RPS4Y1 is knocked out (Figure 6.5I). TNC is closely implicated in promoting increased cell adhesion and migration in asthma. Therefore, we identify that RPS4Y1 has a complex and dynamic role in regulating the transcription and translation of extracellular matrix proteins. In particular, RPS4Y1 appears to mediate cell migration and attachment by producing tenascin-C and interacting with crucial integrin receptors identified in Figures 6.5H & 6.5I.

Figure 6.5: Comparison of gene expression and protein abundance between wildtype and RPS4Y1 KO cell lines. (A) Volcano plot of -$\log_{10}$ (p – value) against $\log_2$fold-change of differentially expressed genes (DEGs) between wildtype and RPS4Y1 KOs after 48-hour incubation at 37°C/ 5% CO_2. Horizontal black lines represented a false discovery rate < 0.05, dotted red vertical lines indicate $\log_2$fold-change > |1.0|. Red dots indicate upregulation, blue dots represent downregulation in RPS4Y1 KOs, and grey dots represent non-significant genes; n = 3. GSVA of up (B) or down (C) DEGs in RPS4Y1 KOs tracked in protein abundance dataset; n = 3. $\log_2$ counts per million (CPM) of *FN1* (D), *COL4A1* (F) and *TNC* (H); n = 3. Fibronectin (E), collagen 4α1 (G) and tenascin-C (I) were measured by ECM ELISA. Higher levels of absorbance (y-axis) indicate a higher level of protein abundance; n = 6 – 9. Data are presented as the arithmetic mean +/- SEM and analysed by unpaired parametric t-test. Statistical significance is represented by *p<0.05 and **p<0.01.

6.3.6 Genes altered in RPS4Y1 knockouts are associated with asthma severity in males

RPS4Y1 has broad regulatory functions affecting inflammatory and fibrotic factors. As such, we conducted GSVA using genes upregulated in RPS4Y1 KOs and tracked their expression in bronchial biopsies from healthy patients and patients with asthma. We observe that genes upregulated in RPS4Y1 KOs are significantly decreased in asthmatic patients (Figure 6.6A). Furth*er,* using a linear regression model, *RPS4Y1* expression positively correlates with FEV1 % predicted in healthy and asthmatic males (Figure 6.6B). That is, increased expression of *RPS4Y1* correlates with better lung function outcomes. We followed up these data by exploring how genes upregulated in RPS4Y1 KOs contribute to FEV_1 % predicted scores using linear regression analysis. We observed that genes upregulated when RPS4Y1 is knocked out do not correlate with any change in FEV_1% predicted in healthy male and female patients (Figure 6.6C & 6.6E) or female asthma patients (Figure 6.6F). However, we report a significant positive correlation in males with asthma (Figure 6.6D). These results indicate that genes usually suppressed by the function of RPS4Y1 are associated with worse lung function outcomes and asthma severity. When these genes have higher levels of expression, patients demonstrate better lung function measurements. As females do not express *RPS4Y1*, we do not expect a significant effect to be observed. However, as *RPS4X* expression does not change in asthma (Figure 6.1F), reduced expression of *RPS4Y1* has functional consequences as recognised through FEV_1 % predicted measurement in asthma.
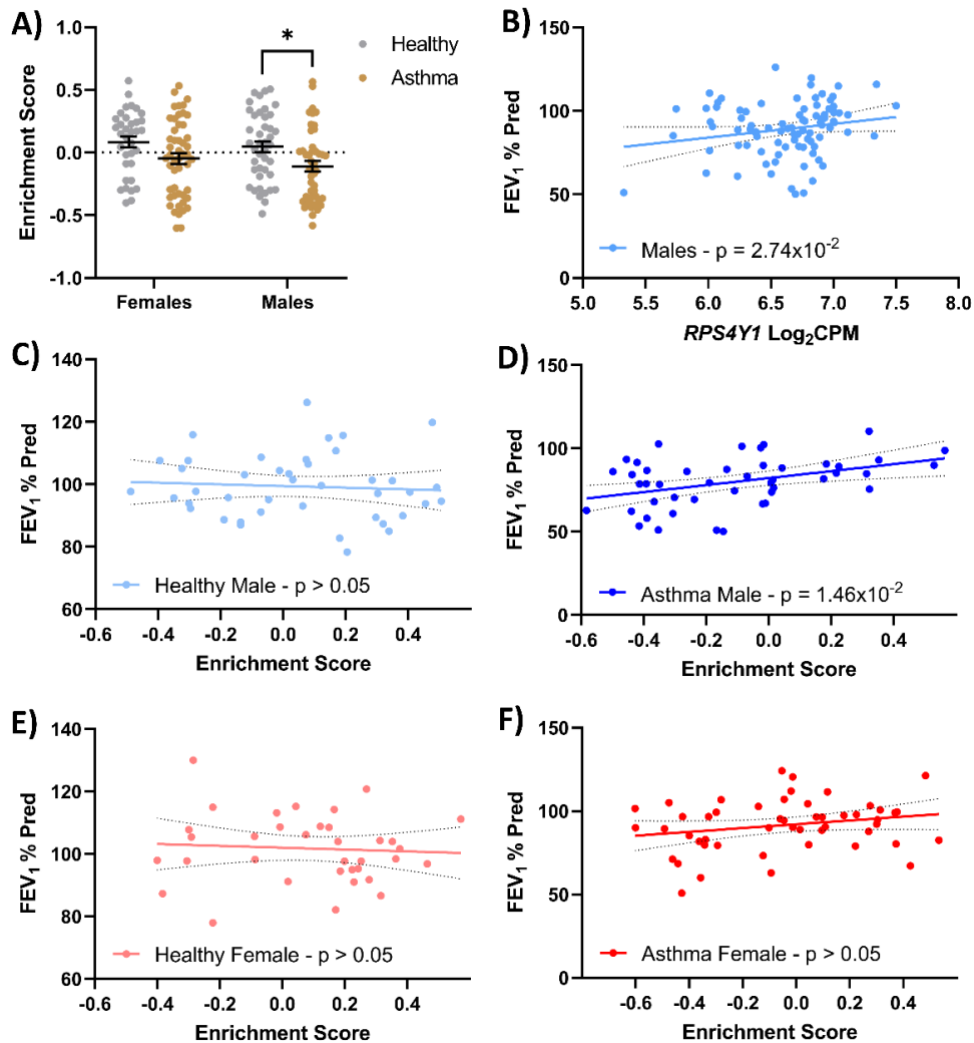
Figure 6.6: *RPS4Y1* **expression and GSVA of upregulated genes in RPS4Y1 KO cells at baseline in healthy and asthmatic patients.** (A) A subset of the top 50 significant upregulated genes in RPS4Y1 KO cells compared to wildtype cells were analysed for how they change in patients with and without asthma. Grey represents healthy patients, and gold represents asthmatic patients. Data are presented as the arithmetic mean +/- standard error of the mean and analysed by two-way ANOVA with Sidak's correction for multiple comparison testing. Statistical significance is represented by *$p < 0.05$ (B) Correlation of $\log_2$ counts per million (CPM) expression of RPS4Y1 against FEV_1 % predicted scores in males (n = 88). GSVA enrichment scores generated in (A) were correlated with the FEV_1 % predicted scored for each patient and stratified by sex and disease. (C) healthy males; n = 42, (D) asthmatic males; n = 46, (E) healthy females; n = 35 and (F) asthmatic females; n = 50. Data are presented with 95% confidence intervals. Statistical significance by linear regression analysis with statistical significance determined at a p-value < 0.05 with correction for smoking pack years.

6.4 Discussion

This study aimed to explore the function of RPS4Y1 in the context of asthma by investigating its contribution to regulating inflammation and fibrotic processes such as cell adhesion and migration. We initially highlight that *RPS4Y1* expression is cell-specific and different to its X-linked counterpart, *RPS4X,* in asthma. Further, *RPS4Y1* expression was lower in male asthma patients than in non-asthmatics, while RPS4X expression did not change for either sex. We show that RPS4Y1 is involved in regulating proinflammatory cytokines CXCL8 and IL6. The combination of transcriptomic and proteomic analyses allowed us to explore potential pathways by which RPS4Y1 may regulate CXCL8 and IL6 production. As such, we uncover that in contrast to CXCL8, whose mRNA expression is altered, IL6 protein translation is increased whilst mRNA expression remains unchanged in the wildtype compared to RPS4Y1 KOs. The apparent selectivity of RPS4Y1 to regulate the translation of protein extended to extracellular matrix proteins that are dysregulated in asthma [98-100], which we hypothesise promotes altered fibrotic processes of cell adhesion and migration. Finally, we generate an RPS4Y1-specific gene signature that only indicates an association with disease severity in asthmatic males. Altogether these data highlight that RPS4Y1 function is altered in asthma and may contribute to unique gene expression profiles between males and females.

Some studies have identified an association between RPS4Y1 expression and asthma [92], through various immune cells [93, 101]. Our study shows RPS4Y1 affects the production of CXCL8 and IL6, which mediate the infiltration and activation of immune cells such as neutrophils and macrophages, driving the proinflammatory response in the airways [102]. A similar finding was reported by Chen et al. [103], where RPS4Y1 knockdown reduced CXCL8 and IL6 production in endothelial cells with high glucose stimulation. In contrast, we show in RPS4Y1 KOs, CXCL8 production is suppressed whilst IL6 is increased. This indicates that the contribution of RPS4Y1 may be cell and stimuli specific.

SPDEF gene expression is significantly reduced in RPS4Y1 KOs at both baseline and after TNFα stimulation, with this change reported to be greater than that in wildtype cells. *SPDEF* is well-known to be increased in airway epithelial cells of asthma patients [104, 105], inducing goblet cell metaplasia and increasing mucous production [106, 107]. Both mucous hypersecretion and increased goblet cell density are associated with asthma severity via increasing airway wall thickness [108]. McKay and Hogg posit that any alteration in airway wall structure may have worse effects on airflow in male children than females [108]. Therefore, dysregulation of *SPDEF* expression by RPS4Y1 may predispose male children to asthma development.

Notably, the immunoregulatory effect at a protein level was not completely reflected at a transcriptional level. CXCL8 mRNA production was reduced, like protein levels, which is driven by increased production of NKRF and activation of the FOXA1 transcription factor pathway. Increased NKRF production has been shown to cause suppression of CXCL8 in A549 cells [57]. Comparatively, FOXA1 knockdown was associated with reduced CXCL8 mRNA levels in two cell models and recovered upon induction of FOXA1 [109]. As such, the combination of these two factors being dysregulated in RPS4Y1 KOs results in decreased suppression of CXCL8. Of significant interest, the expression of *IL6* in KO cells was equal to what was measured in wildtype cells. This indicates that in the absence of RPS4Y1, *IL6* mRNA is preferentially translated to protein. This mechanism is highly complex, with increased IL6 protein being produced to the lack of RPS4Y1 subunit enabling greater affinity with the ribosomal resulting in increased translation, or RPS4Y1 may function through extra-ribosomal means that are poorly defined.

Nonetheless, despite as little as 10% of ribosomes containing RPS4Y1 [83], there is a distinct, specific and novel regulatory function of RPS4Y1 in IL6 production. Females do not express RPS4Y1, and further studies are required to determine whether RPS4X carries similar regulatory functions. However, we hypothesise this would not be the case due to the high expression level of RPS4X in males compared to RPS4Y1. This indicates a significant function for RPS4X supported by most ribosomes containing the X-linked version. Therefore, it is highly likely that RPS4Y1 carries an immunoregulatory function distinct from its X-linked counterpart. This difference may contribute to immunological differences observed been males and females with asthma.

Pathway analysis revealed that RPS4Y1 also functions to regulate cell adhesion and migration. RPS4Y1 knockdown has been associated with increased cell invasion [110]. In particular, integrin-mediated cell adhesion is identified to be enriched in RPS4Y1 KO cells, driven by increased expression of *ITGA4* and decreased expression of *ITGB8*. The relative expression of both integrin subunits is associated with increased adhesion to fibronectin and the promotion of cell migration. The extracellular matrix (ECM) is a highly complex scaffold that is carefully regulated in the lungs to enable robust response in the airways to external stimuli. As such, any dysregulation in the production of ECM proteins may contribute to airway remodelling and worse lung function outcomes in asthma patients. We identify a complex regulatory function of RPS4Y1 for three pertinent proteins in the context of asthma; fibronectin, tenascin-C (TNC) and collagen 4α1. Although RPS4Y1 KOs have increased transcription of fibronectin and collagen 4α1, this does not increase protein production levels.

In contrast, *TNC* gene expression is similar between wildtype and RPS4Y1 KOs, yet KO cells demonstrate significantly increased production of TNC protein. This is unique, as TNC is usually reduced in normal adult tissues [65], yet TNC is increased in the airways of asthmatics [68]. TNC is known to directly contribute to processes of cell motility and migration [66], which may explain the increased rate of wound closure observed in RPS4Y1 KOs compared to wildtype cells. These results reinforce the complex and particular function of RPS4Y1 in promoting and impeding the translation of ECM protein. Nonetheless, these data highlight a strong relationship between RPS4Y1 and ECM proteins that are prominent in asthma, indicating a role for RPS4Y1 in asthma progression.

The functional contribution of RPS4Y1 to the clinical outcomes of patients with asthma is confirmed through our correlation of GSVA enrichment scores and lung function outcomes. Genes upregulated in RPS4Y1 KOs cells at baseline demonstrated a positive correlation with FEV_1 % predicted scores only in male patients with asthma. FEV_1 % predicted describes the volume of air expelled by an individual as a percentage of the average values by healthy individuals [111]. FEV_1 % predicted is a useful clinical tool for determining asthma severity, with scores less than 60% considered severe [111, 112]. Therefore, we have identified a gene signature associated with RPS4Y1 that is only altered in males and correlates with asthma progression. Importantly, this result reinforces that asthma-related disease pathways are distinct between males and females and require deeper investigation.

A multi-omics approach incorporating RNA-sequencing, proteomics and function studies is a significant strength of this investigation. This design allowed for use to identify complex regulatory contributions of RPS4Y1 impacting gene transcription and protein translation. Furthermore, we have identified novel functional effects of RPS4Y1 that directly relate to the hallmark features of asthma – inflammation and fibrosis. We have also identified potential pathways by which RPS4Y1 may mediate these outcomes, but further, more specific and carefully designed studies are required to validate and holistically explore these mechanisms.

Due to the vital role of RPS4X, evidenced by the analysis of the DepMap database and is supported by multiple studies [78, 83, 84], it was not possible to generate and RPS4X KO cell line. Cases of Turner syndrome, where one copy of the X chromosome is missing in females, highlights that only one copy of RPS4X is necessary for survival, although various medical and developmental problem exist [79]. Therefore, due to reduced overall expression of *RPS4X/Y1* in males (Figure 6.1C & 6.1D) indicates that males may carry a greater health susceptibility to disease development. This increased predisposition likely manifests in childhood, before the sex hormone production and activity are increased. *RPS4Y1* expression correlates with FEV_1 % predicted scores, whereas decreased *RPS4Y1* expression correlates

with increased asthma severity. Comparatively, *RPS4X* reports no significant correlation. This potentially indicates distinct functions between the two genes.

Here, we have identified RPS4Y1 regulates hallmark asthma-relevant pathological processes of inflammation, cell adhesion and migration, and mediation of the extracellular matrix. These results implicate RPS4Y1 in the divergent development and progression of asthma between males and females. The use of an inducible RPS4X knockout model or small interfering RNAs may enable the investigation of the function of RPS4X, enabling a comparison to RPS4Y1. This work would enable a deeper understanding of ribosomal mechanisms and specific differences between RPS4Y1 and RPS4X. As such, these deeper investigations of the identified pathways and mechanisms will open the opportunity for identifying new and more effective clinical interventions to improve patient outcomes.

6.5 Conclusion

We show for the first time that RPS4Y1 regulates the protein translation of specific proteins, namely IL6 and ECM factors. We further identify the regulatory transcription factor pathway altered when RPS4Y1 is knocked out, potentially contributing to suppressed gene expression and protein production of CXCL8. Furthermore, we show that RPS4Y1 mediates the expression of integrins subunits α4 and β8, which associate with tenascin-C increasing cell adhesion and migration, which are known to contribute to worse asthma outcomes. We also establish that RPS4Y1 regulates a specific gene signature that is only dysregulated in asthma males, indicating that the imbalance of RPS4Y1 and RPS4X expression between males and females may contribute to sex differences in asthma. Importantly, this study provides more weight to the growing work investigating sex differences in asthma and other prominent respiratory diseases. As such, future studies can build on the findings of the current work to better explore the mechanisms driving sexual dimorphism in disease.

6.6 Supplementary Figures
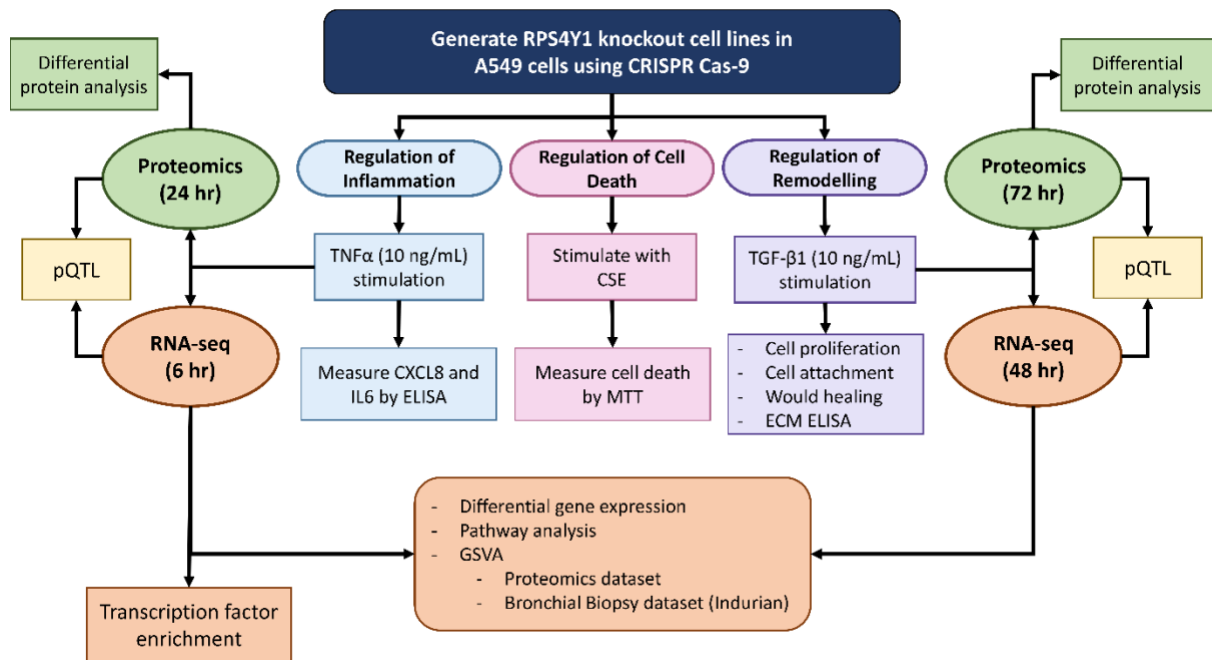
Figure S6.1: Schematic presentation of Chapter 6 study design and techniques used.
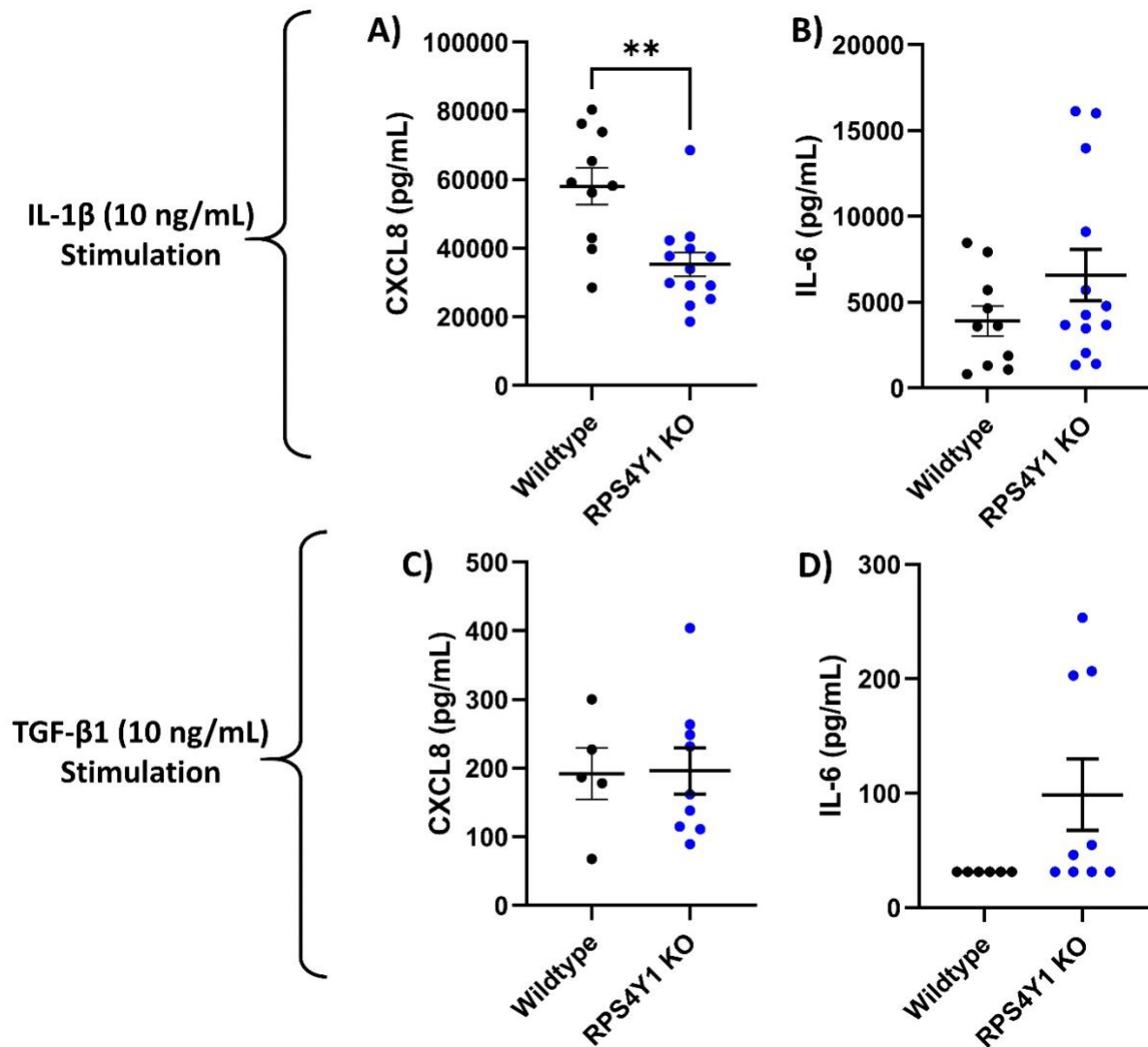
Figure S6.2: Production of CXCL8 (A & C) and IL6 (B & D) from RPS4Y1 knockout and wildtype cells when stimulated with IL-1β (A & B) and TGF-β1 (C & D). CXCL8 and IL6 were measured after 24-hour IL-1β (10 ng/mL) and TGF-β1 (10 ng/mL) stimulation in cell-free supernatant by ELISA. All data are presented as the arithmetic mean +/- SEM. One-way ANOVA statistical analysis with Tukey's correction for multiple comparisons. Statistical significance is indicated by **$p < 0.01$; n = 5 – 13.

Chapter 7 – The histone demethylase, UTY, uniquely contributes to COPD severity

7.1 Introduction

Chronic obstructive pulmonary disease (COPD) is a chronic respiratory disease defined by progressive and irreversible airflow obstruction [113]. As a result, patients experience a decline in lung function, assessed by the forced expiratory volume in one second (FEV_1) and the forced vital capacity (FVC). Clinically COPD is diagnosed when the ratio of these values, FEV_1/FVC, is less than 0.7 post-bronchodilator treatment [114]. After diagnosis, COPD severity is separated into four stages determined by the patient's FEV_1 measurement as a percentage of the predicted FEV_1 of the general population, as outlined by the Global Initiative for COPD (GOLD). COPD is the 3rd leading cause of death worldwide and has a worsening mortality trajectory [115]. Oxidative stress is introduced by noxious stimuli such as those cigarette smoke and biofuel exhaust [116], leading to cellular dysfunction and cell death[117]. An imbalance of regulated cell death (apoptosis) and cell renewal of structural cells contributes to lung destruction instigated by inhalation of cigarette smoke and the development of emphysema [117, 118]. Aberrant cell death correlates with the development of emphysema and worse patient outcomes [119, 120]. The exact molecular cause of COPD remains unknown. Cigarette smoking is the most well-defined risk factor [121]; however, approximately 25 – 45% of individuals with COPD have never smoked [122]. This implies that non-smoking-related factors such as genetic predisposition, exposure to other stimuli such as biomass fuel, and occupational exposures over time also contribute to COPD development. As such, COPD is considered a disease of adult-onset due to a lifetime of exposure to noxious stimuli [123]. Therefore, an intricate interaction between environmental exposures and molecular and genetic factors contributes to COPD development and progression.

Compounding the complexity of COPD is that it has a distinct sexually dimorphic pattern, where adult females are potentially more at risk than males [28, 124]. Historically, males have been reported to have an increased incidence of COPD, although there is growing evidence towards greater susceptibility in females [125, 126]. This pattern is attributed to the normalisation of smoking rates between males and females. Studies indicate that female smokers are 50% more likely to develop COPD and experience increased exacerbation rates than males [124, 127, 128]. Of note, disease pathology also demonstrates a sexually dimorphic pattern. Males present with increased levels of emphysema, with females having more inflammation of the airways [127]. The factors and mechanisms causing this discrepancy in COPD severity and progression between males and females remain poorly understood and

understood. This is highlighted by DeMeo [129], who discusses the critical need to explore multiple facets of COPD to generate better biomarkers and improve patient outcomes.

The sex chromosomes present as a distinct inherent factor distinguishing males and females. Males carry an X and Y chromosome, whilst females express two X chromosomes. Complex epigenetic and molecular mechanisms exist to correct for an imbalance of gene expression, where one female X chromosome is effectively inactivated [130]. However, a subset of genes with Y-chromosome-linked counterparts escapes X-inactivation. Importantly, these genes demonstrate essential genome regulatory functions such as the histone demethylation gene pair KDM6A (UTX) and KDM6C (UTY). Both genes are described to catalyse the removal of trimethyl groups from lysine-27 residues of histone H3 (H327me3), with the demethylation of this site associated with gene activation [131]. Notably, both increased or decreased methylation of H3K27 have been implicated in disease severity and progression [132, 133]. Generally, histone methylation in COPD is poorly understood. Bryd et al. describe a significant decrease of H3K27 trimethylation in COPD bronchiolar epithelium [134], whilst another study identified increased levels of H3K27me3 in COPD patients [135]. Differences in study populations or sampling locations may cause this discrepancy. Anzalone et al. [135] note that an increase in H3K27me3 occurred irrespective of smoking status, despite previous studies identifying H3K27 methylation as sensitive to cigarette smoke [136]. This indicates that COPD-specific changes in the H3K27me3 status may occur after smoking cessation, promoting disease progression.

UTX has an 84% amino acid similarity with UTY [137], with 98% similarity identified across the histone demethylase domain[138]. As a result of this difference, some studies have determined that UTY lacks activity as a demethylase *in vitro* [139, 140]. However, a more recent study by Walport et al. [141] demonstrated that UTY does have an active histone demethylase site, although its activity may be reduced as site-specific and potentially includes non-histone targets. The importance of UTY is supported by studies showing deletion of UTX is fatal in females [138, 142], whilst males without UTX survive [141]. Therefore, some overlapping functions between UTX and UTY exist. In particular, their catalytic target, H3K27me3, is undoubtedly altered in COPD; however, how it relates to cigarette smoking remains unclear.

In this study, we aim to explore the expression and contribution of UTX and UTY to the response to cigarette smoke, regulation of the inflammatory response, cell proliferation and cell death. We identify a unique gene expression pattern for UTY between the nasal and bronchial epithelium of COPD patients, which does not occur for UTX. Using knockout cell lines, we identify unique gene signatures differentially regulated by UTX and UTY, shedding

light on distinct biological pathways altered by either gene. Cumulatively, we generate a signature correlating with COPD disease severity, which appears unique to males. As such, our study uncovers distinct pathways regulated between UTX and UTY and highlights a potential cell survival mechanism mediated by UTY. As females do not express UTY, this differential function may contribute to sex differences in COPD.

7.2 Methods

7.2.1 Analysis of *UTX* and *UTY* expression in non-COPD and COPD patients

The SHERLOCk RNA-seq dataset assessed *UTX* and *UTY* gene expression in paired nasal and bronchial brushings from non-COPD and COPD patients. *UTY* and *UTX* gene expression was evaluated in both nasal and bronchial brushing samples in male and female patients. Log_2 counts per million (CPM) were analysed for correlation with FEV_1/FVC scores, using linear regression models accounting for age and pack years. The data were stratified by disease status according to GOLD stage of disease definitions [143]. The patient demographics of this study are located in section 3.16.5.

7.2.2 Generation of CRISPR Cas9 knockout cell lines

Using CRISPR-Cas9, three unique knockout (KO) cell lines were established in A549 cells for UTX and UTY, with UTX/UTY double-knockouts (DKO) also generated, as described in section 3.2. Double-knockout cells were generated using an established UTY KO cell by completing the CRISPR-Cas9 process targeting a UTX exon. A schematic representation of the overall study design and analyses is included in the supplement as Supplementary Figure S7.1

7.2.3 Cell culture and treatments

All generated cell lines were maintained in DMEM growth medium (DMEM supplemented with 10% (v/v) FBS and 1% (v/v) antibiotic/antimycotic). A detailed description of cell maintenance technique and cell treatments are located in section 3.1.

7.2.4 Western blot

UTX and UTY knockouts were confirmed by western blot analysis. Anti-human UTX (D3Q1l, #33S10S, Cell Signalling) and anti-human UTY (#PA5-68440, Invitrogen) were used to detect and visualise protein bands. H3K27me3 abundance was also measured in KO samples using anti-human H3K27me3 (323, #61018, Biosearch). Detailed methodology is located in section 3.19.

7.2.5 Measurement of CXCL8 and IL6 protein secretion

Wildtype, UTX KO, UTY KO and DKO cell-free supernatants were collected 24 hours post-stimulation with TNFα (10 ng/mL). The concentration of CXCL8 and IL6 was quantified by ELISA and was completed as described in section 3.3.

7.2.6 RNA-sequencing

Whole-cell RNA extracts from wildtype, UTX KO, UTY KO and DKO cells were collected 48 hours post TGF-β1 (10 ng/mL) stimulation and processed as described in section 3.10 and section 3.11.

7.2.7 Differential gene expression analysis

Differential gene expression analysis was completed using the *Dseq2* package in R. Wildtype, UTX KO, UTY KO and DKO cell lines were compared 48 hours after serum starvation at baseline (no treatment) and 48-hour TGF-β1 stimulation. A detailed description of differential gene expression (DGE) analysis is located in section 3.12.

7.2.8 LC-MS/MS proteomics analysis

Proteomics analysis was completed for wildtype, UTX KO, UTY KO and DKO cell lines as described in section 3.17. Protein lysate samples were collected 72 hours after serum starvation at baseline (no treatment) and 72-hour TGF-β1 (10 ng/mL) stimulation.

7.2.9 Gene set variation analysis (GSVA)

GSVA was completed as described in section 3.14. This analysis used the protein abundance dataset generated by LC-MS/MS to compare the transcriptome and the proteome of UTX KO, UTY KO and DKO cell lines by tracking the production of DEGs to their corresponding proteins in the LC-MS/MS analysis dataset.

GSVA was also completed on RNA-sequencing data from nasal brushing collected in the OLiVIA study. Enrichment scores for DEGs identified for UTX KO, UTY KO and DKO cell lines were analysed in never, ex and current smoker patients. Patient demographics are described in section 3.16.4.

The relationship between GSVA enrichment scores for DEGs associated with each knockout cell line with COPD stage and FEV_1/FVC for patients recruited for the SHERLOCk study was also assessed. This analysis was further stratified by sex. Patient demographics are described in section 3.16.5.

7.2.10 Analysis of biological pathways enriched in knockout cell lines

Pathway analysis was completed using the g: Profiler online tool for UTX KO, UTY KO and DKO cell lines, as described in section 3.15.

7.2.11 Cell proliferation

The cellular proliferation rate of wildtype, UTX KO, UTY KO and DKO cells was assessed by manual cell counting after 96-hour incubation in growth medium as described in section 3.9.

7.2.12 Cigarette smoke extract (CSE) generation

Marlboro Red standard cigarettes (Philip Morris) were used to prepare CSE, as described in section 3.6.

7.2.13 Cell viability assay

The cellular viability of wildtype, UTX KO, UTY KO and DKO cells after 24 and 48-hour CSE exposure was assessed using the MTT assay as described in section 3.7.

7.2.14 pQTL validation

Gene expression correlation with protein abundance levels from UTX KO, UTY KO, and wildtype was analysed using pQTL for wildtype, ZFX KO and ZFY KO cells as described in section 3.18.

7.2.15 CSE-induced cell death in primary human airway smooth muscle cells

The use of human tissue for this study was approved by the Ethics Review Committee of St Vincent's Hospital Sydney and the University of Technology Sydney Human Research Ethics Committee (UTS HREC: ETH16-0507; St Vincent's Hospital HREC/15/SVH/351). Written and informed consent was obtained from all volunteers or next of kin.

Airways were isolated from the lung and small airways (less than 2 mm) in patients undergoing lung transplantation or resection for thoracic malignancies, as previously described[144]. Primary human airway smooth muscle (hASM) cells were grown in a T175 flask with DMEM growth medium. The cells were incubated at 37°C/ 5% CO_2. At 80% confluency, cells were seeded at $3.4x10^4$ cells/mL, as described in section 3.1, in a 6-well cell culture plate for the following experiments. Once confluency was achieved, primary hASM cells were exposed to cigarette smoke extract following the methods described in section 3.6. However, only one cigarette was bubbled through 25 mL of DMEM for 2 min, constituting 100% CSE. hASM cells were incubated with 15% and 20% CSE at 37°C/5% CO_2 for 24 hours. Cell death was measured using two methods, the lactase dehydrogenase assay [144] and using mitotracker green [145], which have been previously described. An Accuri C6 flow cytometer was used to enumerate the live cells containing mitochondria stained with mitotracker green.

Primary hASM cells were treated with MitoQ, to assess the impact of blocking mitochondrial reactive oxygen species (ROS) production. Briefly, hASM cells were pre-treated with MitoQ (50 nM) for 18 hours. Cells were then exposed to CSE at different concentrations post-ligand washout. The clinical features of the patients used for this work are summarised in Table 7.1.

Table 7.1: Clinical summary of Non-COPD and COPD patients. Data are presented as the arithmetic mean +/- standard deviation.

	Non-COPD	COPD
n	7	7
Age, yr	66 (+/- 5.35)	55 (+/- 7.19)
Smoking History	> 40 pack years	> 40 pack years
FEV$_1$/FVC	> 0.7	< 0.4
Surgery	Resection	Transplant

FEV$_1$ = Forced expiratory volume in one second; FVC = forced vital capacity.

7.3 Results

7.3.1 UTY but not UTX gene expression correlates with COPD severity

To examine whether *UTY* expression equivocates for the double dosage of *UTX*, we analysed the expression of both genes using paired nasal and bronchial brushing samples collected in the SHERLOCk study. Figures 7.1A & 7.1G demonstrate that the sum of *UTX* and *UTY* expression in males is greater than the total expression of *UTX* in females in both the nasal and bronchial samples. This difference is caused by significantly greater expression of *UTY* compared to *UTX* in males, where *UTX* expression in males is approximately 50% of that in females. Therefore, we observe an imbalance in both *UTX* and combined *UTX* and *UTY* expression between males and females.

Considering this imbalance, we aimed to explore whether *UTX* and *UTY* expression changes in COPD. In the nasal epithelium, *UTX* expression is unchanged in males and females (Figure 7.1B), but *UTY* is significantly increased in males (Figure 7.1C). This is directly opposite to what is observed in the bronchial epithelium, where *UTX* expression decreases in females (Figure 7.1H) and *UTY* expression decreases in males with COPD (Figure 7.1I). COPD is primarily a disease of the lower airways and respiratory tract. Nonetheless, it is important to recognise gene regulation differences in the nasal epithelium distinct from the bronchial epithelium. We explored this phenomenon further, reporting that changes in gene expression directly correlate with the severity of the disease (Figure 7.1F and 7.1L). Linear regression analysis of *UTX* and *UTY* gene expression with FEV_1/FVC ratio revealed that only *UTY* correlated with lung function outcomes in COPD patients (Figure 7.1D and 7.1J). However, no significant association was reported for *UTX*. As such, *UTY* expression reflects the severity of COPD progression in patients. *UTX* expression in both males and females is reduced in the bronchial epithelium compared to the nasal irrespective of COPD diagnosis (Figure 7.1M & 7.1N). However, UTY expression remains similar between the nasal and bronchus in non-COPD patients yet shows a significant decrease in the bronchial epithelium in COPD patients (Figure 7.1O). This indicates that *UTY* is dynamically regulated in response to COPD diagnosis. These data demonstrate an essential and complex contribution of *UTY* to the progression and severity of COPD.
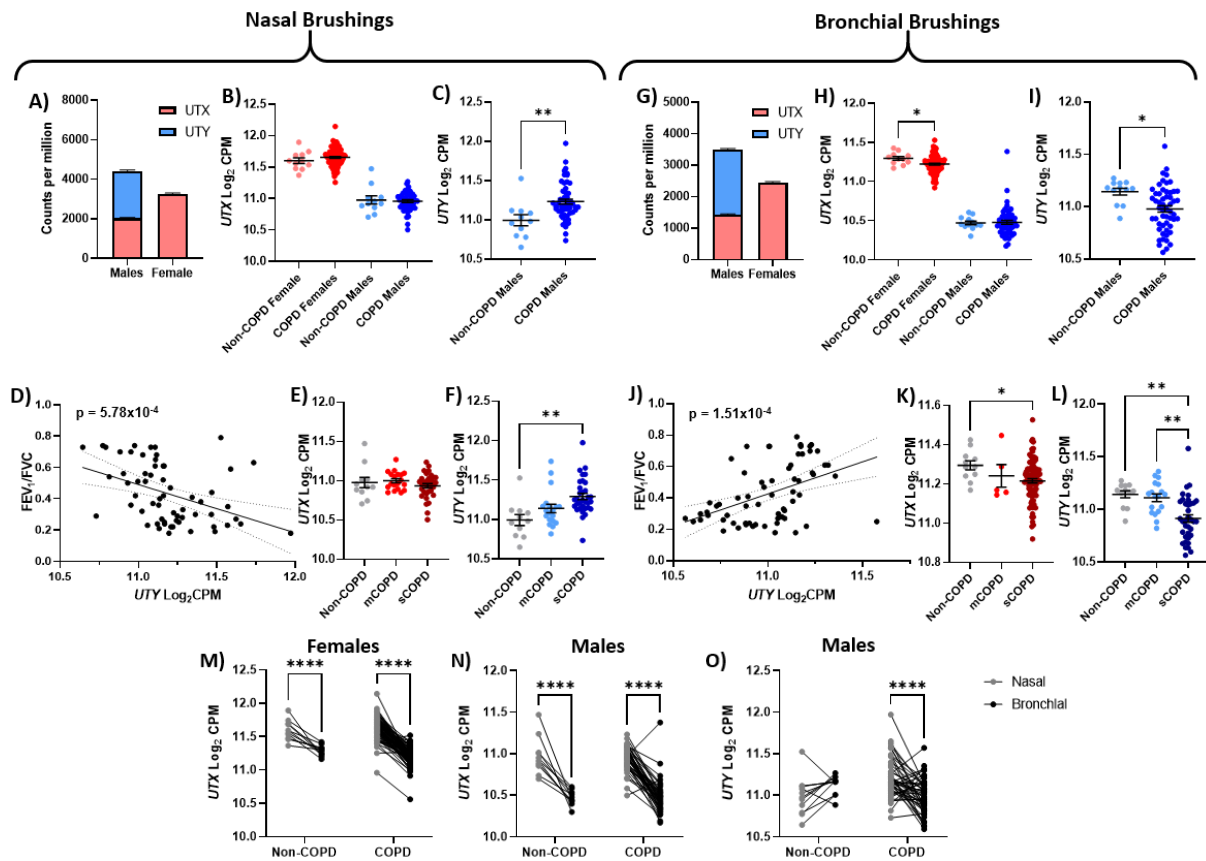
Figure 7.1: Gene expression of *UTX* and *UTY* in nasal brushings (A – F) and bronchial brushings (G – L) from COPD patients. (A & G) Average gene expression of *UTX* (blue) and *UTY* (pink) in males and females in counts per million (CPM). (B, C, H & I) Gene expression of *UTX* and *UTY* stratified by COPD diagnosis and sex. (D & J) *UTY* log$_2$CPM gene expression correlated with FEV$_1$/FVC ratio in males with COPD; data were analysed using a linear regression model. (E, F, K & L) *UTX* and *UTY* log$_2$CPM gene expression stratified by COPD disease stage. (M – O) *UTX* and *UTY* log$_2$CPM gene expression in paired nasal (grey) and bronchial (black) brushing samples, stratified by disease diagnosis. All data (except D, J, M – O) presented as the arithmetic mean +/- SEM and analysed by one-way ANOVA with Tukey posthoc test. (M – O) analysed by two-way ANOVA with Sidak correction for multiple comparisons. n = 101 females/89 males nasal brushings and 85 females/88 males bronchial brushings. All UTY analyses only included male patients. Statistical significance was determined at p < 0.05 and signified by *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

7.3.2 Double-knockout of UTX and UTY alters cell death mechanisms in response to cigarette smoke extract

It is essential to explore the overall contribution of UTX and UTY to the pathological processes involved in COPD. As such, we investigated the function of both genes in inflammation, cell death and proliferation by generating A549 knockout cell lines using CRISPR-Cas9. Three different cell lines were generated, UTX knockout, UTY knockout and a UTX/UTY double-knockout (DKO). Confirmation of gene knockout was completed using western blotting of H3K27me3, the target histone site for UTX and UTY (Figure 7.2A), and RNA-sequencing (Figure 7.2B & 7.2C). Genome sequencing confirmed the introduction of stop codons or indels, leading to a truncated protein sequence (Supplementary Figure S7.2). Interestingly, UTX expression indicates a trend towards decreased expression in UTY KOs (Figure 7.2B). Comparatively, UTY gene expression in UTX KOs is similar to wildtype cells (Figure 7.2C). There is no significant change from wildtype levels for both UTX and UTY expression in double-knockout cell lines. This highlights that a complex regulatory network exists between UTX and UTY. Both genes are required for normal function, but some compensation and equivalent function exist in single knockouts.

The functional characterisation of the knockout cell lines supports this. No significant difference in the regulation of CXCL8 and IL6 was observed across all cell lines (Figure 7.2D & 7.2E). A trend towards reduced proliferation (Figure 7.2F) and doubling time (hours) was found in UTX (24 hours) and UTY (25 hours) KOs, but a significantly slower rate of proliferation in DKO (33 hours) compared wildtype cells (23 hours).

Cigarette smoking is a major contributory factor to COPD development and progression. As such, Figures 7.2G and 7.2H explore the different contributions of UTX and UTY in regulating cell death induced by cigarette smoke extract (CSE). After 24-hour exposure, UTY KOs reported higher survival levels than wildtype cells after exposure to 50% and 75% CSE (Figure 7.2G), but this difference does not exist after 48 hours. No change in survival is seen for UTX KOs. A significant increase in survival is seen for DKO cell lines. After 24 hours, more than a 75% survival rate is seen for all CSE concentrations (Figure 7.2G). After 48 hours of exposure to 50% CSE, 80% cell survival is recorded for DKO cells compared to no treatment, whilst the death rate is similar to wildtype at higher CSE concentrations.
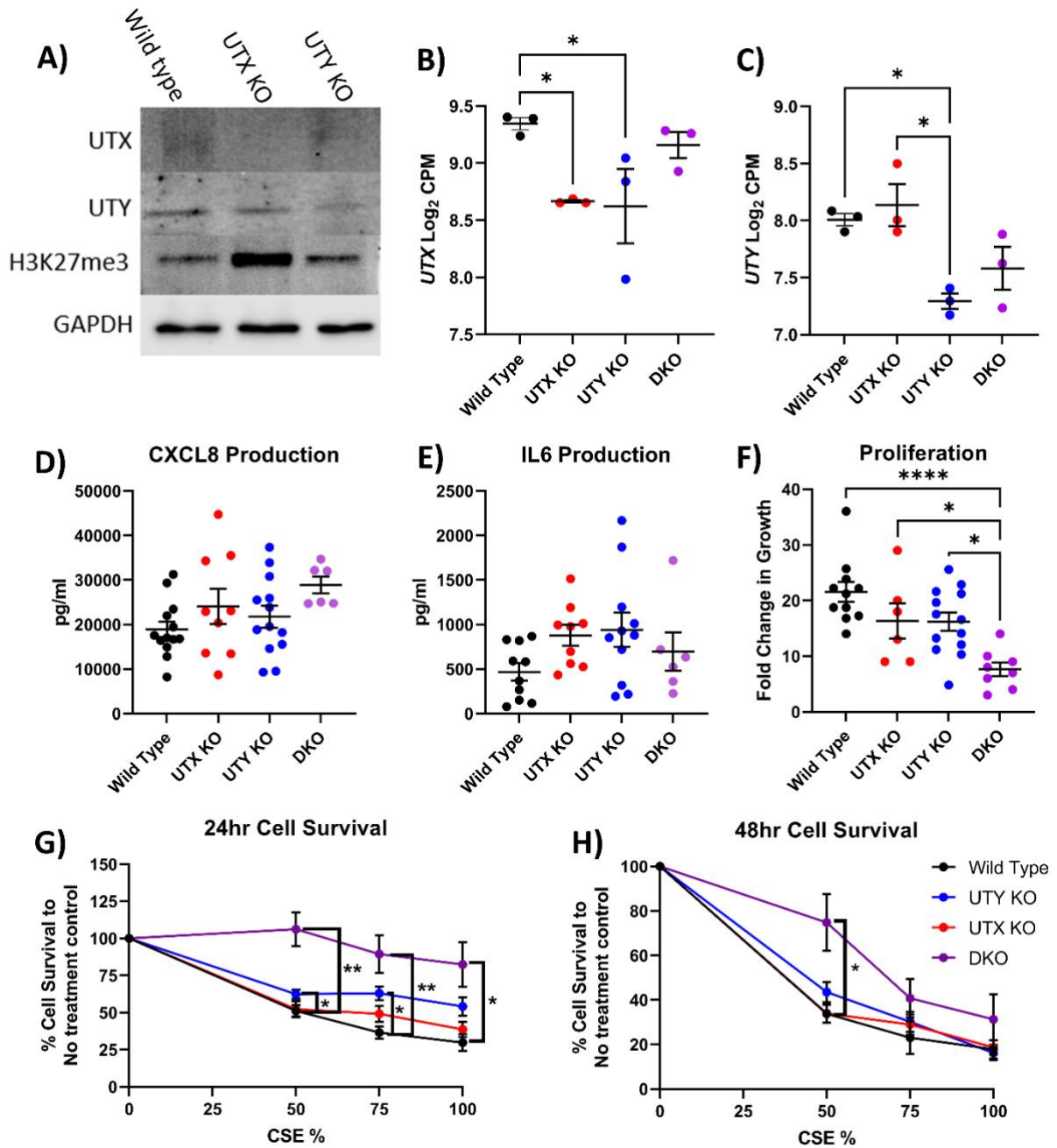
Figure 7.2: Changes in regulation of inflammation, proliferation and cell survival in knockout cell lines. (A) Western blot analysis of H3K27me3 in wildtype, UTX KO and UTY KO cell lines. (B & C) $\log_2$ counts per million (CPM) gene expression for UTX (B) and UTY (C) across all cell lines. (B – F) the x-axis indicates the genotype of the cell line. CXCL8 (D) and IL6 (E) were measured after 24-hour TNFα (10 ng/mL) stimulation in cell-free supernatant by ELISA. (F) Cell counting measured proliferation after 96 hours of incubation at 37°C/5% CO2 in growth medium. Percentage cell survival compared to control was analysed using MTT assay 24 (G) and 48 hours (H) after cigarette smoke exposure (CSE). All data are presented as mean +/- SEM. (B – F) One-way ANOVA statistical analysis was conducted using Tukey's multiple comparison corrections; $*p<0.05$, $****p<0.0001$, n=6-12. (D – E) Two-way ANOVA with two-stage step-up Benjamini, Kreiger and Yekutieli multiple comparisons correction method used; $*p<0.05$, $**p<0.01$, n=8-9. Black dots represent wildtype cell lines, red represents UTX knockout cells, blue represents UTY knockout cells, and purple represents DKO.

7.3.3 Knockout of UTX primarily contributes to the expression profile of the double-knockout cell line

We completed RNA-sequencing analysis to explore genes and pathways that are differentially regulated in UTX, UTY and double-knockout cell lines. Differential gene expression analysis was conducted without stimulation to observe differences at baseline compared to wildtype cells. UTX KOs reported 1215 differentially expressed genes (DEGs), UTY KOs had 1435 DEGs and double-knockout cells presented 1725 DEGs (Figure 7.3A – 7.3C).

We compared the single knockout cell lines at baseline to explore the distinct functions of UTX and UTY. However, we removed the DEGs when either single KO cell lines were compared to wildtype cells to investigate how they affect different pathways. This analysis revealed genes exclusively regulated between UTX and UTY, highlighting genes regulated explicitly by either gene. This analysis revealed 76 differentially expressed genes distinctly regulated by either UTX or UTY (Figure 7.3D). This finding emphasises that UTX and UTY have similar targets, but a distinct subset of gene targets does exist, potentially contributing to the functional differences reported in Figure 7.1. We completed a similar double-knockout analysis method but subtracted significant differentially expressed genes by UTX and UTY single knockouts at baseline from the double-knockout comparison to wildtype cells. Subsequently, 787 genes were reported to be uniquely differentially expressed in the DKO cell lines (Figure 7.3E). This analysis reveals a significant subset of differentially regulated genes to compensate for the loss of both UTX and UTY. These genes are not dysregulated when either UTX or UTY are individually knocked out compared to wildtype cells. As such, this gene list represents genes that are similarly regulated by both UTX and UTY. Table 7.2 summarises pathway analysis of the top 50 genes up or downregulated in the above analyses. Histidine catabolism is positively enriched in UTY knockout cells with mitochondrion-related processes downregulated. These pathways may contribute to the differential response of UTY to cigarette smoke exposure.

The heatmaps (Figure 7.3F & 7.3G) demonstrate how differentially regulated genes in DKOs (Figure 7.3E) and unique to UTX and UTY (Figure 7.3G) are altered across all cell lines. Interestingly, Figure 3G shows genes upregulated in UTX KOs are similarly upregulated in DKOs. This finding indicates that the majority of change in responsiveness observed in DKOs is due to the loss of UTX, which has a uniquely distinct effect on gene expression outside of the function of UTX. The gene *HSP90AB1* was significantly upregulated across all knockout cell lines compared to wildtype cells (Figure 7.3H). This gene encodes the heat shock protein 90 alpha B1 (HSP90ab1) which is closely associated with apoptotic processes and the mitochondrial function[146]. *CHORDC1* expression was increased in UTY and DKO knockout cell lines compared to wildtype cells, with expression from DKO cells also significantly

increasing compared to wildtype cells (Figure 7.3I). This gene functions as a co-chaperone with HSP90ab1 across various cellular processes. This data indicates that HSP90ab1 increased expression paired with increased expression of *CHORDC1* promotes increased cell survival to CSE, as depicted in Figure 7.2H.
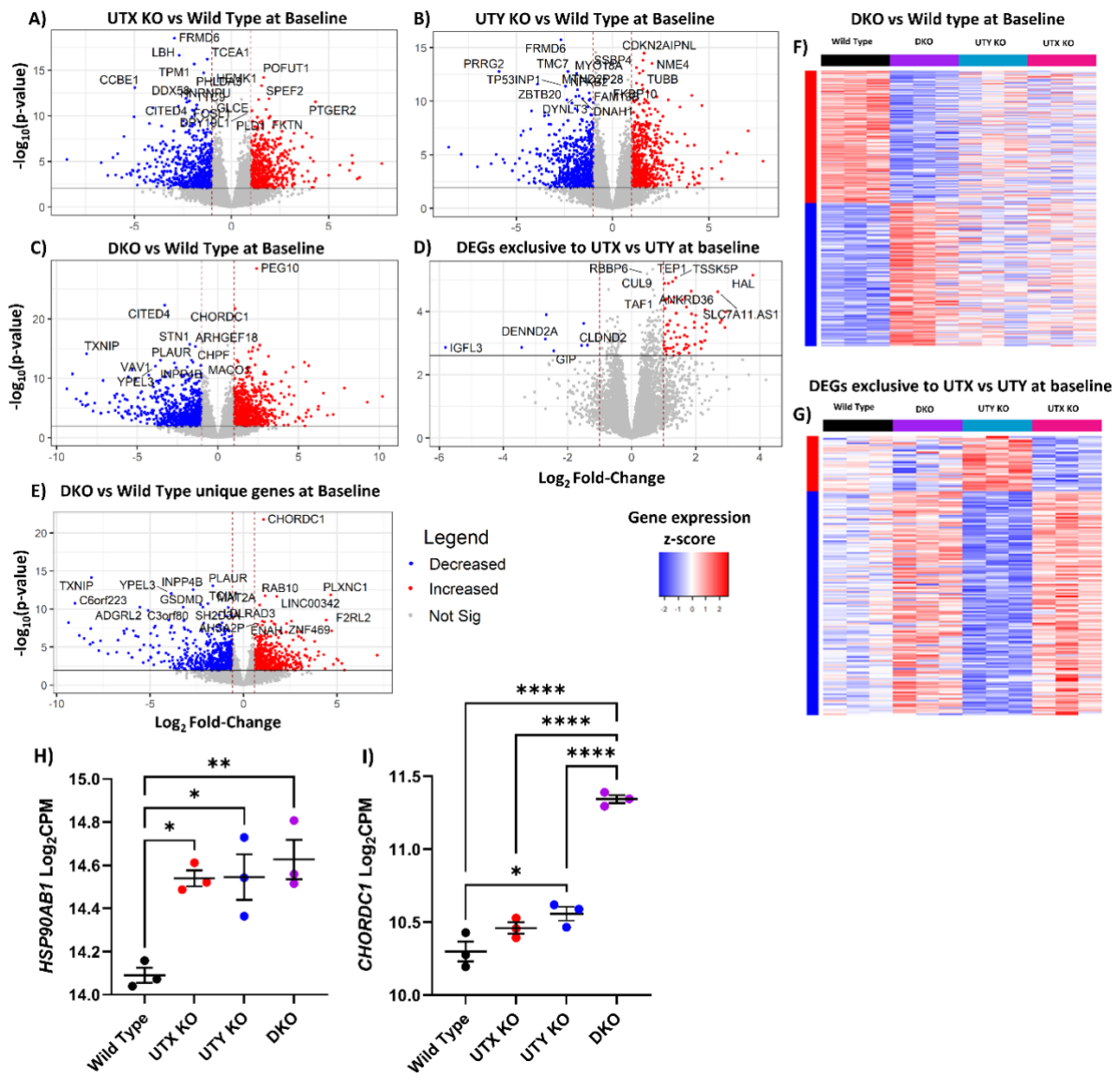
Figure 7.3: Differential gene expression analysis of UTX and UTY knockout cell lines.
(A – D) Volcano plots of -$\log_{10}$(p-value) against $\log_2$ Fold-change in gene expression for UTX (A), UTY (B) and double-knockout (C) cell lines against wildtype cells at baseline. (D) presents genes differentially expressed between UTX and UTY knockout cell lines, which are not altered when either individual KO is compared to wildtype cells. (E) presents genes that are differentially expressed between DKO and wildtype cells that are not also changed when UTX or UTY are knocked out alone. Statistical significance was determined at a false discovery rate < 0.05 (horizontal line) and $\log_2$FC > |1.0| (vertical lines). (E & F) Supervised heatmap of differentially expressed genes in the double-knockout cell line (E) and the DEGs unique to UTX and UTY (F) at baseline. Genes with increased expression are coloured red, and decreased genes are coloured blue. The genes are tracked for wildtype (purple), double-knockout (yellow), UTY KO (orange) and UTX KO (green) cells. $\log_2$ counts per million (CPM) gene expression for *HSP90AB1* (H) and *CHORDC1* (I) presented as the arithmetic mean +/- SEM. Data is analysed by one-way ANOVA with Tukey's correction for multiple comparisons. Statistical significance is represented by *p<0.05, **p<0.01 and ****p<0.0001; n=3 for all analyses.

Table 7.2: Summary of g: profiler pathway analysis using genes up and downregulated genes across all knockout cell lines at baseline.

Increased in UTX KO at baseline		Increased in UTY KO at baseline		Increased in DKO at baseline	
Enriched Pathway	**FDR**	**Enriched Pathway**	**FDR**	**Enriched Pathway**	**FDR**
Cell surface receptor signalling	4.5×10^{-2}	Histidine catabolic process	1.8×10^{-2}	Cell surface receptor pathway	2.7×10^{-2}
Cell death	2.1×10^{-2}	Histidine metabolic process	2.9×10^{-2}	Basement membrane	3.3×10^{-4}
Decreased in UTX KO at baseline		**Decreased in UTY KO at baseline**		**Decreased in DKO at baseline**	
Enriched Pathway	**FDR**	**Enriched Pathway**	**FDR**	**Enriched Pathway**	**FDR**
Superoxide-generating NADPH oxidase activator activity	4.8×10^{-2}	Mitochondrion	6.0×10^{-3}	Cellular response to stimulus	1.1×10^{-2}
Secretory vesicle	2.4×10^{-2}	EGR-EP300 complex	4.9×10^{-2}	Signal transduction	1.3×10^{-2}

FDR = false discovery rate; NADPH = nicotinamide adenine dinucleotide phosphate; EGR = early growth response 1

7.3.4 UTX and UTY can partially compensate for the loss of the other

Differential gene expression analysis indicated an ability for UTX and UTY to compensate for the loss of the other gene. We used GSVA analysis to explore this concept in more detail across all cell lines. Genes upregulated when UTX or UTY is knocked out are upregulated across all knockout cell lines (Figure 7.4A & 7.4B). Similarly, genes downregulated in UTX and UTY KOs are also downregulated across all knockout cell lines (Figure 7.4E & 7.4F). Notably, genes are downregulated at a greater level when UTX and UTY are knocked out, indicating these genes function as more potent gene activators than repressors. This is confirmed by other studies showing H3K27me3 as a repressive mark [147]. This notion is further supported when investigating genes up (Figure 7.4C) and downregulated (Figure 7.4G) by DKO cells. Importantly, these results show that genes dysregulated in DKO cells are similarly altered in the single knockout cell lines, although to a lesser extent. Figures 7.4D & 7.4H highlight that the transcriptome of UTX KO cells is more similar to DKO cells, with UTY KOs demonstrating a distinctly opposite regulation of the same gene set. Importantly, this reinforces that UTX and UTY have different effects on gene regulation. These GSVA results show UTX and UTY cannot compensate when the other is absent. Interestingly, we highlight that a unique and distinct set of genes are upregulated when UTX and UTY are knocked out, which are not altered compared to wildtype cells in single knockout cells.
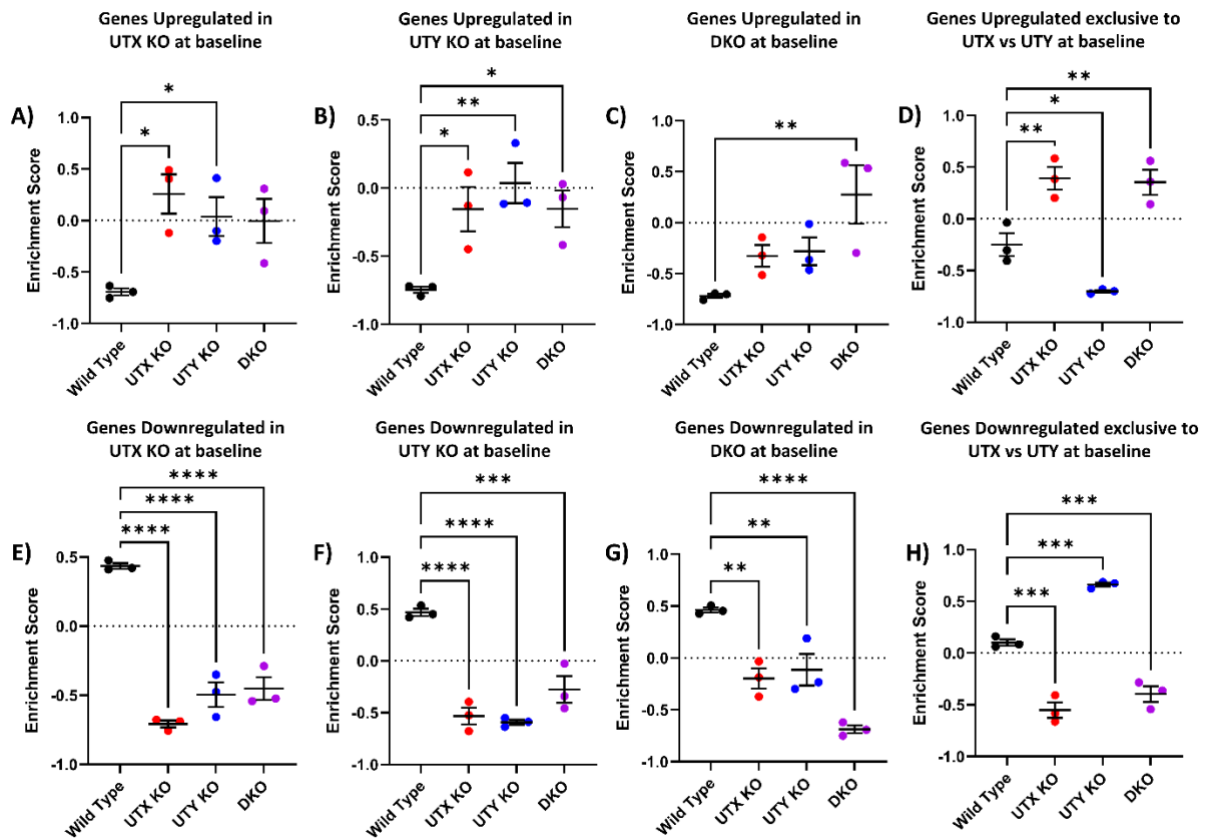
Figure 7.4: GSVA analysis of differentially expressed genes across all cell lines. Genes upregulated (A – D) or downregulated (E – H) in UTX, UTY and double-knockout cell lines to wildtype cells at baseline. Data are presented as arithmetic mean +/- SEM and analysed using one-way ANOVA with Tukey correction for multiple comparisons. Statistical significance is indicated by *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. n = 3.

7.3.5 Changes in gene expression are reflected at a protein level

We have established unique responses between all knockout cell lines at a transcriptomic level. However, it is necessary to complete peptide-centric proteomic analysis to confirm whether these changes in gene expression are functionally apparent at a protein level. We used LC/MS-MS analysis with GSVA to determine whether differentially expressed genes were tracked through a protein abundance dataset. Three proteins were significantly downregulated in UTX KOs compared to wildtype cells at baseline (*PLS1*, *ANXA13* and *MCU*). The same three genes were downregulated in UTY KOs with *SLC9A3R1* also decreased (Figure 7.5A & 7.5B). More proteins were differentially expressed in the DKO cell lines, with 41 proteins increased, and 196 proteins decreased, compared to wildtype cells. Both *PLS1* and *ANXA13* were the two most significant differentially expressed proteins across all cell lines. Figure 7.5D shows that DKO cells have a unique expression profile at a protein level compared to UTY and UTX KOs, which are relatively similar. The profile of this protein heatmap shows a more significant difference between DKOs and the single knockout cell line than in the gene expression heatmap (Figure 7.3F). This observation, combined with the significantly increased number of differentially expressed proteins in DKO compared to single knockouts, indicates that UTX and UTY affect translational processes and factors.

Across all three differential protein analyses, PLS1 (Plastin 1), MCU (Mitochondrial calcium uniporter), and ANXA13 (Annexin A13) are consistently significantly downregulated (FDR < 0.05; $\log_2$FC < -1). We report these proteins to demonstrate equally reduced protein abundance in all knockout cell lines compared to wildtype cells (Figure 7.5E – 7.5G). This indicates that UTX and UTY are vital regulators of these proteins, with the presence of both UTX and UTY together necessary for normal expression.

GSVA analysis of significant differentially expressed genes tracked in the protein dataset demonstrated that changes in gene expression and similarly reflected at the protein level (Supplementary Figure S7.3A – S7.3F). This analysis reinforces that double-knockout of UTX and UTY has a compounding effect on gene expression. Genes up or downregulated in DKO are also increased or decreased in single knockout cell lines, although less than what is reported when both genes are knocked out (Supplementary Figure S7.3C & S7.3F). As such, UTX and UTY have overlapping targets, where the presence of one can partially compensate for the loss of the other.

A) UTX KO vs Wild Type at Baseline

B) UTY KO vs Wild Type at Baseline

C) DKO vs Wild Type at Baseline

Legend • Decreased • Increased • Not Sig

D) DKO vs Wild Type at Baseline

Protein Expression Z-score

PLS1

E)

MCU

F)

ANXA13

G)
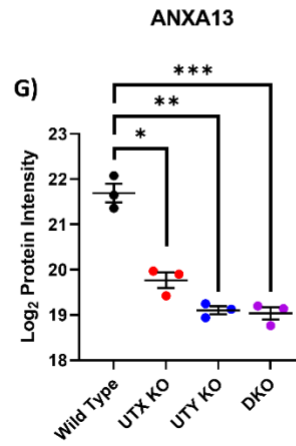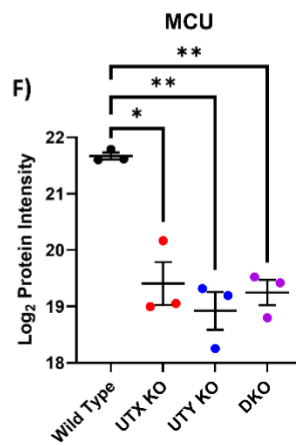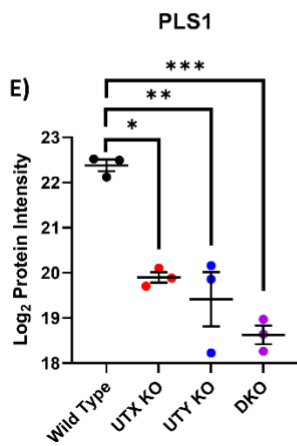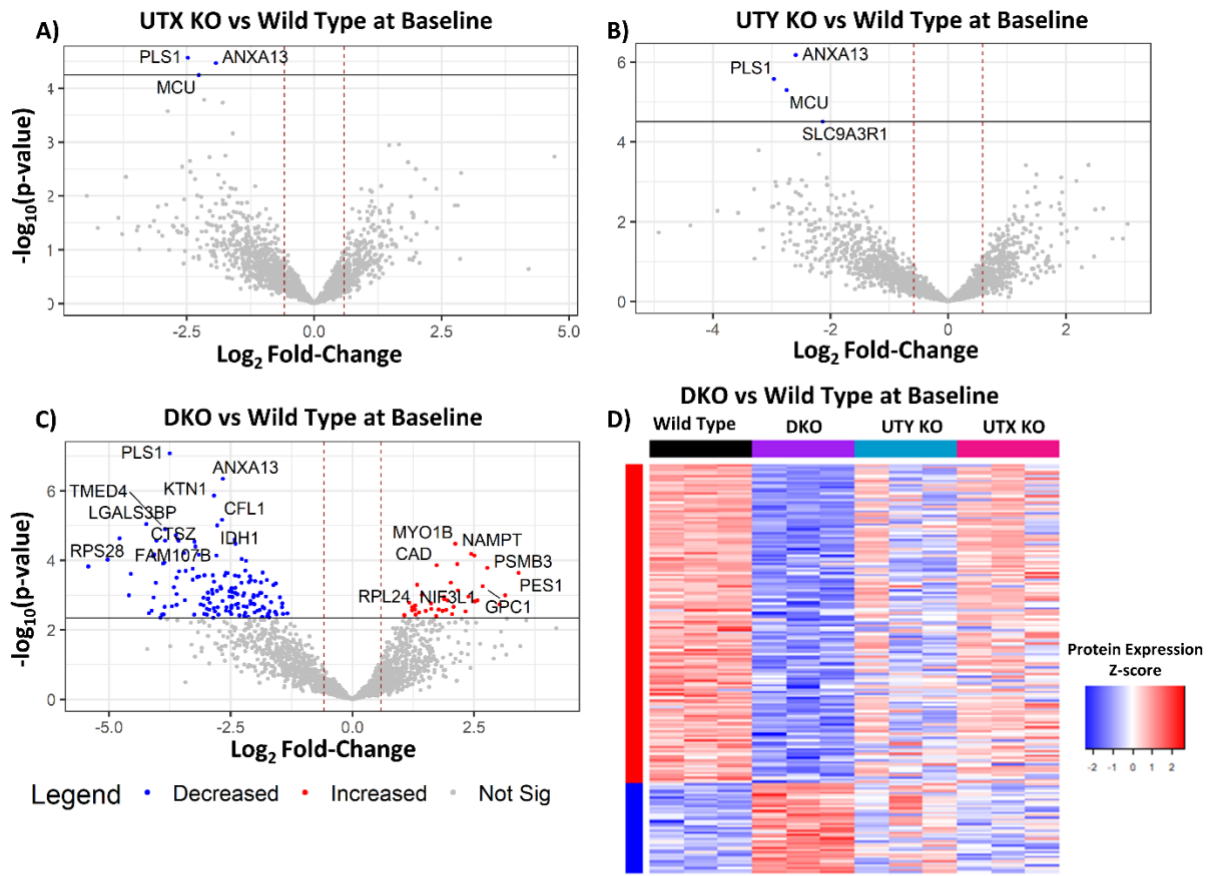
Figure 7.5: GSVA of differentially expressed genes in proteomics dataset. (A – C) Volcano plots of -$\log_{10}$(P-value) against $\log_2$ fold change in protein abundance for UTX (A), UTY (B) and double-knockout (C) cell lines compared to wildtype cells at baseline. The solid horizontal black line indicates a false discovery rate (FDR) of 0.05. Dotted red vertical lines represent $\log_2$ fold change |1.0|. Blue dots represent statistically significant proteins that are downregulated; red dots represent statistically significant proteins that are upregulated. (D) Supervised heatmap of differentially expressed genes in the double-knockout cells line across all cell lines used. Genes with increased expression are coloured red, and decreased genes are coloured blue. The genes are tracked for wildtype (purple), double-knockout (yellow), UTY KO (orange) and UTX KO (green) cells. $\log_2$ Protein intensity measured by LC-MS/MS of PLS1 (E), MCU (F) and ANXA13 (G) in all cell lines at baseline. Data is analysed by differential protein analysis with statistical significance, an FDR < 0.05, and Benjamini-Hochberg correction for multiple comparison testing. (E - G) Data are presented as the arithmetic mean +/- SEM. Statistical significance is indicated by *FDR<0.05, **$p<0.01$ and ***$p<0.001$. n = 3 for all analyses.

7.3.6 Differentially regulated genes in UTX/UTY double-knockout are altered with smoking status

Figures 7.2G & 7.2H demonstrated significantly increased survival rates in DKO cells compared to wildtype, with a slightly increased survival rate in UTY KOs. As such, exploring how differentially expressed genes in the knockout cell lines are altered in patients with different smoking histories was pertinent. We conducted GSVA analysis of DEGs identified in our knockout cell lines within a cohort of never, ex and current smoking patients[20]. The results show that genes upregulated or downregulated in UTX (Figure 7.6A & 7.6E) and UTY (Figure 7.6B & 7.6F) knockout cell lines are not altered with smoking status. Genes upregulated in DKO cells also show no change (Figure 7.6C). However, downregulated genes indicate increased expression in ex-smoker nasal epithelium compared to never-smoker patients (Figure 7.6G).

When the refined list of differentially expressed genes unique to the DKO cell line is used, we report distinct changes in gene expression depending on smoking status. These genes are only differentially expressed in DKO cells and not altered in the single knockout cell line. Genes that are only upregulated in DKOs compared to wildtype cells are increased in ex-smokers compared to both never and current smoker individuals, with no difference reported between never and current smokers (Figure 7.6D). The reverse is observed for genes that are only downregulated in DKO cells. These genes are also downregulated in the nasal epithelium of ex-smokers but are not significantly changed in never or current-smoker patients (Figure 7.6H). These findings indicate that UTX and UTY function together to produce an adaptive response to cigarette smoke when the stimulus is removed. This highlights a potential mechanism for epigenetic modifications in response to cigarette smoke exposure in the respiratory tract.
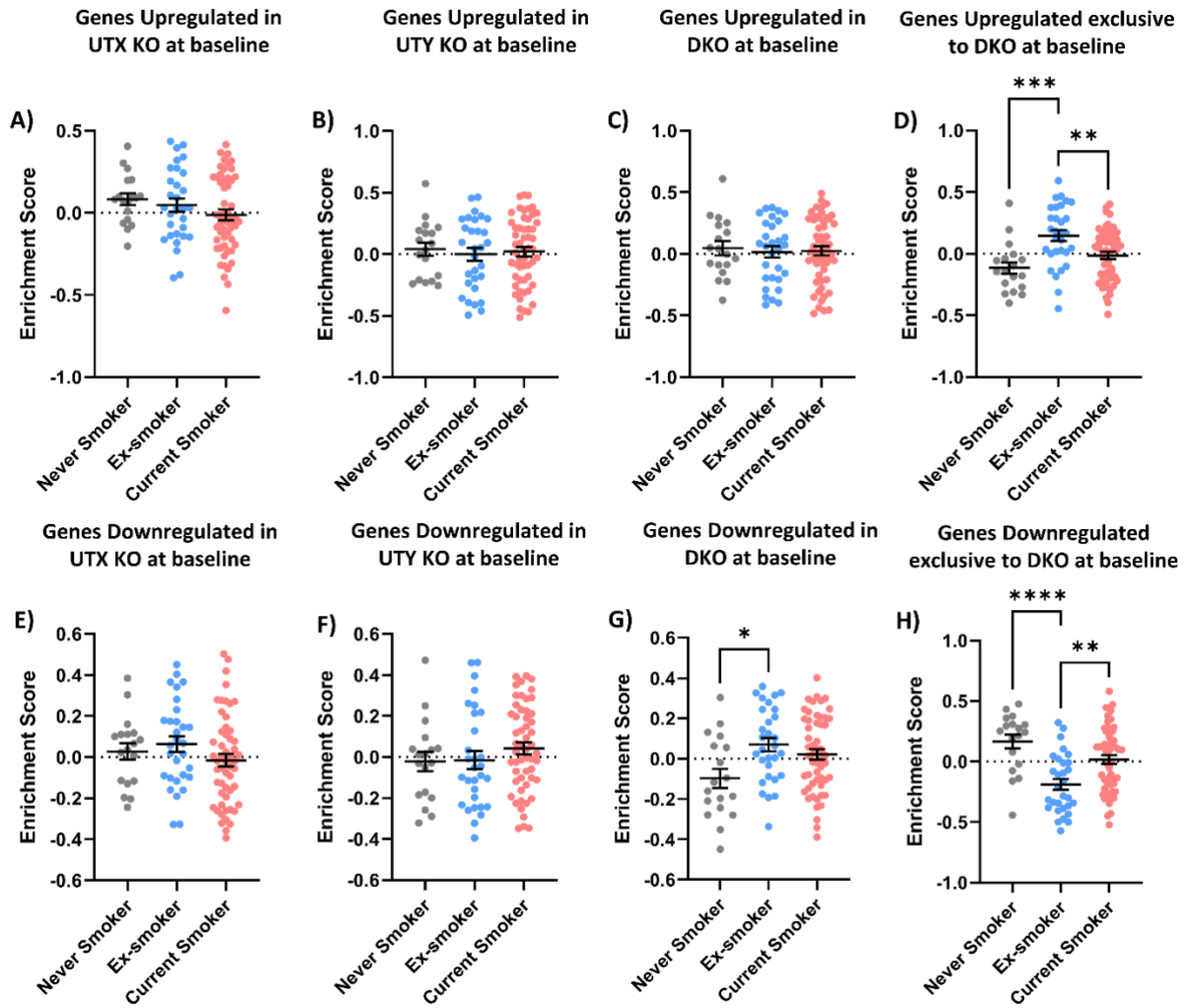
Figure 7.6: Enrichment of differentially expressed genes in knockout cell lines in nasal epithelium of patients with different smoking statuses. Grey dots = never smokers, blue dots = ex-smokers and red dots = current smokers. A subset of the top 50 significantly increased (A – D) and decreased (E – H) genes in all knockout cell lines. Data are presented as the arithmetic mean +/- SEM. One-way ANOVA completed statistical analysis with Tukey's post hoc test. Statistical significance is represented by *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001; n = 36 never smokers/ 30 ex-smokers/ 36 current smokers.

7.3.7 Mitochondrial and H3K27 functions are altered in COPD affecting cell death mechanisms

A link between UTX and UTY and cigarette smoking exists (the primary factor contributing to COPD development). We highlight increased cell survival in UTY KO cells compared to wildtype cells. Dysregulated cell death is an essential pathological feature contributing to the destruction of lung tissue and alveolar structures. We aimed to investigate differences in cell death responses to cigarette smoke extract in primary cells isolated from the airways of COPD and non-COPD patients. Figure 7.7A shows increased cell death after 20% CSE exposure in non-COPD patient cells compared to baseline, whilst no difference is reported for COPD cells. This finding is replicated using mitotracker green (Figure 7.7B). This methodology highlights a distinct loss of mitochondria in non-COPD cells after CSE exposure, which does not occur in COPD-derived cells. These results establish that the regulation of cell death in COPD is aberrant, leading to increased cell survival rates, with mitochondrial-related pathways potentially involved. Subsequently, cells were treated with MitoQ – a targeted antioxidant compound that reduces the production of reactive oxygen species (ROS) from mitochondria [148]. Non-COPD patient-derived cells demonstrate increased survival in response to 20% CSE exposure with mitoQ treatment (Figure 7.7C), whilst no change in cell survival occurred in COPD patient-derived cells. Therefore, mitochondrial-related cell death mechanisms are aberrant in COPD. Importantly, all patients in this study are heavy smokers, meaning that difference in survival rates is due to the presence of COPD and is independent of a lifetime smoking effect.

Increased cell survival in COPD-derived cells reflects the increased survival we observed in the DKO cell lines in response to CSE (Figure 7.2G & 7.2H). This pattern indicates that genes regulated by UTX and UTY may be altered in COPD patients, causing increased cell survivability to noxious stimuli. Genes upregulated in DKO cells compared to wildtype are decreased in the nasal epithelium of COPD patients (Figure 7E). This decrease is predominant in severe COPD patients, with mild COPD individuals demonstrating no difference compared to non-COPD patients (Figure 7.7F). In the bronchial epithelium, genes downregulated in DKO cells are increased in COPD patients (Figure 7.7G), with this increase seen only in severe COPD patients (Figure 7.7H). Therefore, we have identified a gene expression signal regulated by UTX and UTY, reflecting the stage of COPD progression.

The FEV_1/FVC ratio is the primary measurement for the severity of COPD. Therefore, we used linear regression analysis to assess the relationship between lung function and GSVA enrichment score from our knockout cell lines. Considering the distinct patterns of COPD susceptibility and severity between males and females combined with the unique expression pattern of UTY (Figure 7.1), we investigated the association between DEGs in UTY KOs and

the severity of COPD. Genes upregulated in DKO cells at baseline positively correlated with lung function scores for both sexes in the nasal epithelium (Figure 7.7I). In contrast, genes downregulated in DKO cells are negatively associated with lung function in the bronchial epithelium (Figure 7.7J). Next, we analysed genes that are regulated explicitly between UTX and UTY. Genes relatively downregulated in UTY KO cells indicate a negative correlation with lung function in males (Figure 7.7K). When UTY is absent, these genes have reduced expression, which is associated with worse lung function outcomes only in male patients. Similarly, when genes that are relatively upregulated when UTY is absent are increased in male patients, the patients indicate worse FEV_1/FVC ratios (Figure 7.7L). Importantly, these correlations highlight that males with COPD demonstrate greater responsiveness to gene expression changes, affecting lung function.
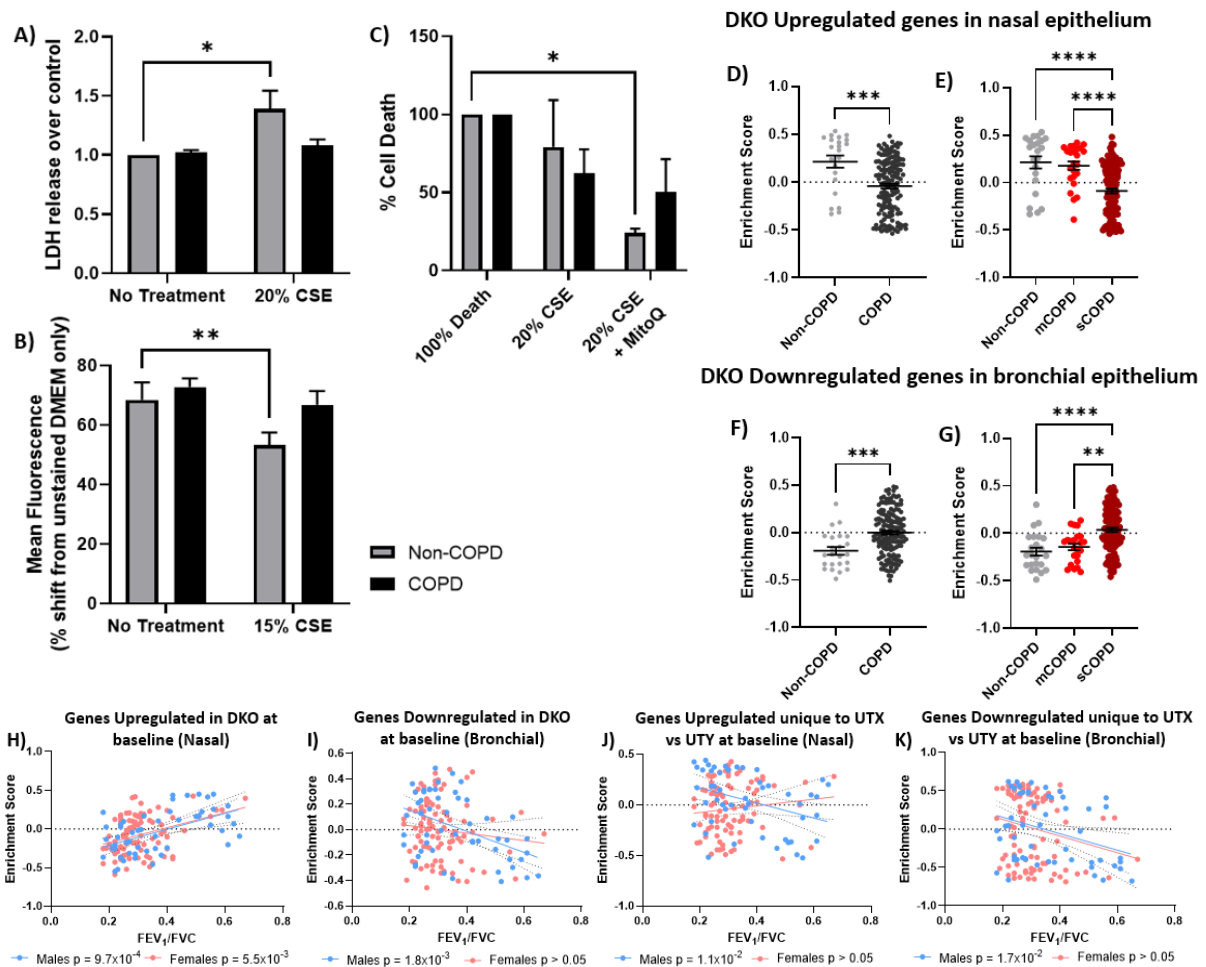
Figure 7.7: Comparison of cell death in non-COPD (grey) and COPD (black) hASM cells (A – C) and comparison of GSVA enrichment scores of differentially expressed genes in double-knockout cells with COPD severity (D – K). (A & C) Measurement of cell death using LDH assay after exposure to 20% cigarette smoke extract between non-COPD (grey) and COPD (black) patients derived hASM cells. (B) Analysis of cell death using mitotracker green dye methodology. Data are presented as mean +/- SEM and analysed by a mixed-effects model with Sidak correction for multiple comparisons; n = 3 non-COPD & 7 COPD. GSVA enrichment scores were generated using the top 50 significant differentially expressed genes in the double-knockout cell line at baseline in non-COPD and COPD patient nasal (D, E, H & J) and bronchial (F, G, I & K) epithelium. Analysis was stratified by disease stage (E & G). Data are presented as the arithmetic mean +/- SEM and analysed by parametric t-test (D & F) or one-way ANOVA with Tukey correction for multiple comparisons (E & G); n = 23 non-COPD/ 24 mCOPD/ 125 sCOPD. Statistical significance is represented by *p<0.05, **p<0.01, ***p<0.001 and ****p<0.0001. (H – K) A multiple linear regression model assessed the correlation between GSVA enrichment score and FEV_1/FVC in COPD patients. Statistical significance was determined at p<0.05; n = 51 males/ 91 females.

7.4 Discussion

We identify a unique expression pattern for *UTY* in COPD patients, where *UTY* is decreased in the nasal epithelium but increased in the bronchial epithelium. This change of expression correlates with disease severity, with the change of expression increasing or decreasing with the stage of COPD for the patient. As such, we highlight a potential contribution or dysregulation of *UTY* in a disease state which may contribute to COPD patient outcomes. Using knockout cell lines, we demonstrate that when UTX and UTY are both knocked out, cell proliferation dramatically decreases and cell survival to noxious stimuli increases significantly. We identify some ability for UTX and UTY to compensate for the loss of the other gene using single knockout cell lines. However, we also identify a list of genes exclusively differentially regulated between UTX and UTY, which we postulate mostly comprises gene targets beyond demethylase functions [138]. We recognise a complex interaction between UTX and UTY function using independent clinical datasets with both the smoking response and COPD development.

The distinct change in the expression of *UTY* in the nasal and bronchial epithelium highlights a complex interplay between the gene and COPD. This observation arguably contradicts the 'field of injury' concept where similar common molecular responses to stimuli are observed throughout the respiratory tract [149]. Although, this pattern may be a protective mechanism of the body in preparation for future insults. The GSVA analysis in patients with different smoking histories revealed an alteration in gene expression patterns observed only in ex-smokers. This trend indicates that the body attempts to form protective changes upon removing the stimuli through epigenetic mechanisms. As COPD is a disease of the lower respiratory tract, affecting bronchioles and alveoli [150, 151]. Therefore, the observed discrepancy in *UTY* expression between the nasal and bronchial epithelium may indicate an interaction between the upper and lower respiratory tracts in COPD.

UTX is described to have significantly more histone demethylase activity than UTY [138, 141, 152], which is confirmed through our western blot analysis. Despite this divergence in the active catalytic domain, multiple studies have proposed redundancy in the functions of UTX and UTY [147, 153], indicating demethylase-independent functions for both genes [154]. Our data support this notion, as we observe a moderate response in single knockout cell lines, whereas the effect is significantly exaggerated in the double-knockout cell line. However, there are exceptions also reported. DKO cells demonstrate significantly increased survival to cigarette smoke extract, and UTY KO cell lines also indicate slightly increased survival. However, no change to wildtype is reported for UTX KO cells. The loss of Y-chromosome-linked genes with smoking and subsequent increased risk of aberrant regulation of cell death is well reported [155-157]. Our data indicate a significant contribution of UTX maintaining and

supporting correct cell death mechanisms. In particular, heat shock protein 90 (HSP90) isoform *HSP90ab1* and *CHORDC1* have been linked with COPD, where dysfunction of HSP90 is associated with decreased risk of COPD development and apoptotic mechanisms involved in hyperoxic cell death [158-160]. We identify *HSP90ab1* is upregulated in all single and double-knockout cell lines, however, *CHORDC1* expression is also increased in UTY and DKOs with no change reported in UTX KOs. This pattern and magnitude of the change in expression directly reflect the altered pattern of cell survival to CSE. Therefore, we propose that UTX and UTY both regulate the expression of HSP90. However, the increased expression of the co-chaperone molecule *CHORDC1* [158] results in a dysregulation of apoptotic processes.

A distinct interaction between cigarette smoking and dysregulation of cell death and proliferation and the pathogenesis is described in the literature [117]. A differential response between non-COPD and COPD-derived cells to CSE has been reported [9]. We show cells isolated from COPD patients have increased survival rates compared to cells derived from non-COPD patients. This pattern reflects our observed response by UTY and DKO cell lines to CSE. Mitotracker Green is a reliable reporter of mitochondrial mass in live cells [161], with mitochondrial level functioning as a marker of apoptotic cell fate. This combined with MitoQ, a mitochondrial-specific antioxidant [162], protecting only non-COPD patient cells from CSE-induced cell death, implicates a dysregulation of mitochondria promoting persistent survival from COPD cells. Mitochondrion-related biological pathways were negatively enriched in UTY KO cells, and the HSP90 proteins interact closely with the mitochondria to regulate cell death [163]. Importantly, patients selected for this analysis were matched for smoking history, indicating that dysregulation of mitochondria is a function specific to COPD diagnosis and not smoking. Collectively, these data indicate that UTX and UTY may function at the nexus of mitochondrial regulation of cell death through HSP90 and CHORDC1. However, future studies are needed to explore and identify whether UTX and UTY are contributing to this mechanism via demethylase or non-demethylase activity.

We identify that UTY potentially contributes more to sex differences in COPD outcomes, although UTX stops greater aberrant pathological processes from occurring. This discrepancy is apparent through the gene signature we generated via GSVA analysis in COPD patients. Gene upregulated or downregulated in DKO cell lines at baseline indicated a significant correlation with COPD severity through FEV_1/FVC scores. Notably, this correlation was predominantly reported in male patients, namely in the bronchial epithelium (the primary site of COPD pathology). This apparent stability of gene expression in females may be due to double dosage UTX from two X-chromosomes.

This study builds upon previous work characterising the function of UTX and UTY. We highlight some redundancy allowing both genes to compensate for losing the other. But also demonstrate that both genes are necessary for correct cell cycle regulation. Our use of both single knockout and double-knockout cell lines allowed for greater depth of analysis and identification of novel functions of UTX and UTY. However, following up on many of the potential mechanisms and pathways identified remained beyond the scope of the current analysis. Nonetheless, this body of work provides a strong basis for future studies to explore the contribution of UTX and UTY in greater detail. Supporting analysis of differential gene expression after CSE exposure may provide more insight into functional differences between the cell lines. Although, the observed baseline differences offer a valid perspective of the pathways basally altered by these genes. Cumulatively these data provide exciting insight and momentum for future studies to continue investigating how UTX and UTY affect sex differences and the progression of COPD.

7.5 Conclusion

In summary, we identify a novel expression profile for *UTY* in COPD throughout the respiratory tract and confirm and expand on the current literature on the functions of UTX and UTY. We show that UTX and UTY, to an extent, can compensate for the loss of the other, but our DKO model reveals more about their combined regulatory role in cell death and the cell cycle. These pathologically relevant processes are shown to be prevalent in COPD, linking dysregulation of UTX but primarily UTY to sex differences in COPD progression and severity. The results of this study provide future investigations with greater scope and potential gene targets. Building upon this knowledge will help find and understand mechanisms causing sex differences in disease, leading to the development of new, more effective treatments to improve patient outcomes.

7.6 Supplementary Figures
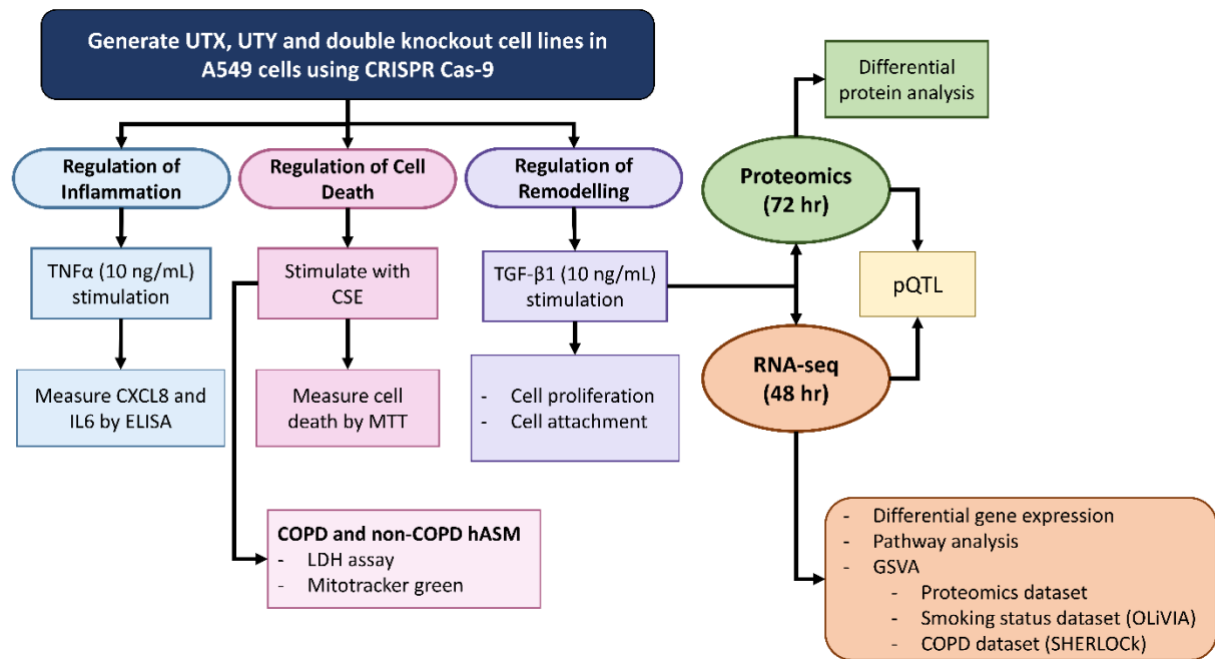
Figure S7.1: Schematic presentation of Chapter 7 study design and techniques used.
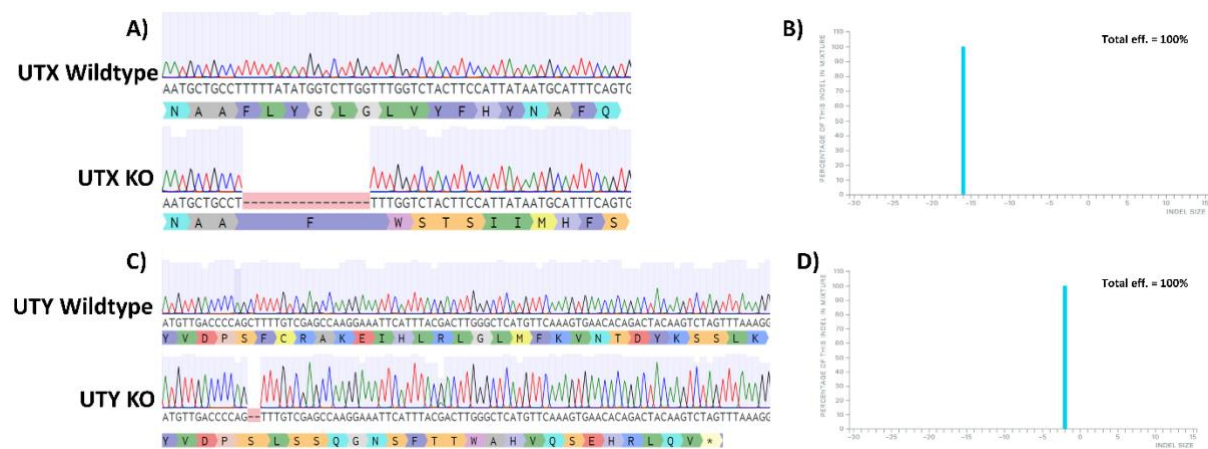
Figure S7.2: Dendrogram and indel graphic of representative of knockout cell lines. Dendrograms of UTX (A) and UTY (C) sequences were generated using Benchling.com and aligned to a wildtype sequence. A deleted base is highlighted in red with a dash (-); 16 bases have been deleted in the representative UTX KO (A), and two bases were deleted in the represented UTY KO sequence. The translated amino acid sequence is presented below each respective dendrogram with an asterisk (*) indicating a stop codon. (B & D) Total eff. = the percentage of the sample sequences that indicate the mutation shown in the dendrograms.
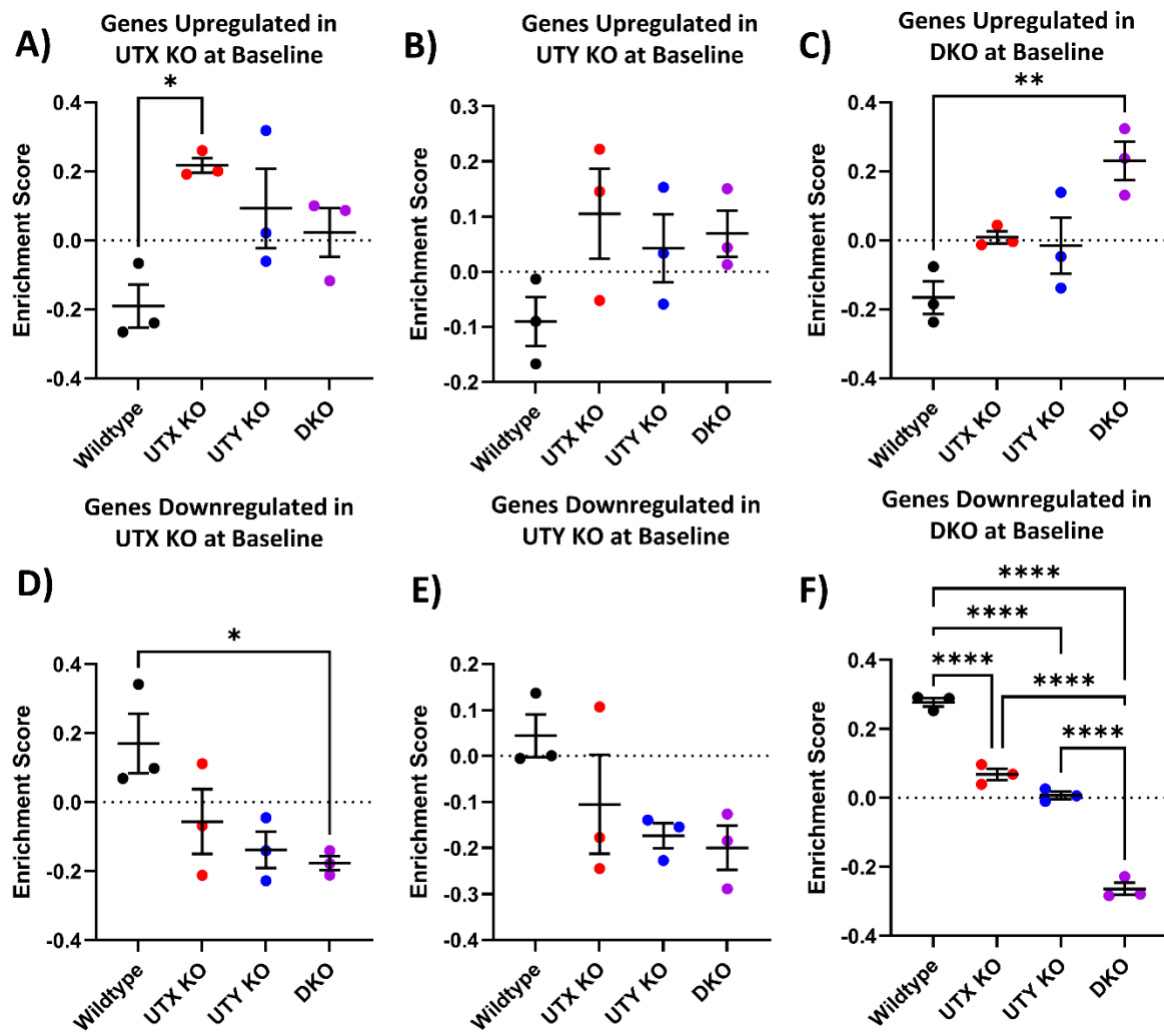
Figure S7.3: GSVA analysis of differentially expressed genes in the proteomics dataset. Genes upregulated (A – C) or downregulated (D – F) in UTX, UTY and double-knockout cell lines to wildtype cells at baseline tracked for their corresponding protein abundance in the proteomics dataset. Data are presented as the arithmetic mean +/- SEM and analysed using one-way ANOVA with Tukey correction for multiple comparisons. Statistical significance is indicated by *p<0.05, **p<0.01, ****p<0.0001. n = 3.

Chapter 8 - General summary, discussion and future directions

Sexual dimorphism refers to differences between males and females at the organ, cellular and molecular levels. This observable divergence between the sexes is critical for establishing biological differences and enabling sexual reproduction [24]. Sexual dimorphism as a phenomenon has been largely understudied and subsequently misunderstood [164]. In particular, many studies have overlooked the impact of biological sex as a determining factor in disease, often considering sex as a covariate or confounding factor in epidemiological studies. This oversight has limited our understanding of differences in disease patterns and processes between males and females. In recent years sex as a critical biological factor has garnered increased attention, motivated by a statement by the National Institute of Health (NIH) [165]. This call highlighted that biology must be considered at all levels of research, from the design of experiments, analysis of data and reporting findings from animal and human studies. This sentiment is reflected throughout multiple research and reviews articles that have delved into the differences between males and females [28, 30, 124, 166]. A key finding noted by these studies is that sex differences exist across multiple diseases, but an individual's response to treatment may also be related to the biological sex[30, 167]. An essential and prudent goal of this is to develop novel, more effective and personalised medical interventions to improve patient health outcomes.

A detailed review of the literature, as presented in **Chapter 1**, highlights that sex differences are prevalent in disease susceptibility, progression and prognosis. This pattern exists across multiple organ systems, with the more susceptible sex varying depending on the disease with the organ [28]. For example, adult males are more likely to suffer idiopathic pulmonary fibrosis [168], whilst adult females with asthma demonstrate increased rates and worse outcomes with asthma [32]. As such, a distinct yet highly complex interaction between biological sex and disease exists.

Asthma and chronic obstructive pulmonary disease (COPD) are two prevalent chronic respiratory diseases. Australia has one of the highest rates of asthma in the world, with 11.2% [169] and COPD presents as the fifth leading cause of death in the country [170]. As such, these conditions significantly burden the Australian healthcare system and population and present as pertinent diseases for further research. Furthermore, both these conditions present a unique pattern of sex differences, making them ideal candidate diseases for the focus of this thesis. Young males demonstrate an increased incidence of asthma compared to young females (12.1% vs 7.9%) [87], although this pattern flips post-puberty towards increased asthma diagnosis in females (13.9%) compared to males (9.6%). COPD indicates greater

severity in females [171]. This may be driven by a response to cigarette smoking, where females show worse lung function than males despite similar smoking rates [128].

Importantly, sex differences in asthma and COPD are observable at a molecular level, leading to differences in pathological processes. For example, immune cell populations between males and females with asthma are distinct and different [172], males experience more significant alveolar destruction in COPD compared to more small airway inflammation and fibrosis in females [127]. Despite the considerable burden of asthma and COPD on society, the exact cause and disease mechanisms for both conditions remain poorly understood. We hypothesise that by exploring the mechanisms promoting sexually dimorphic pathophysiological processes, a deeper holistic understanding of the diseases will occur. This will shed light on fundamental pathologic features, enabling the development of novel treatments for patients.

Chapter 2 presents data that shows that sex differences in the regulation of proinflammatory cytokines CXCL8 and IL6 are observable *in vitro* [173]. This data revealed that irrespective of the disease diagnosis, primary pulmonary fibroblasts isolated from female patients produced more IL6 than male cells when stimulated with TNFα. No change in CXCL8 production was reported between the sexes. Notably, differential IL6 production was observed from cells removed from the human body and hence, removed from the impact of the sex hormones. Therefore, we propose the observed sex difference in inflammation was driven by an intrinsic factor that differs between male and female cells. This conclusion provided evidence that sex chromosome-linked factors may contribute to this effect.

Human males carry one X chromosome and one Y chromosome, whilst females carry two copies of the X chromosomes. To account for this discrepancy, one X-chromosome in females is inactivated [174]. Approximately 15-25% of X-chromosome genes escape inactivation (XCi), with the expression levels of these genes attributed to phenotypic variability between the sexes[175]. As a result, females present with a 'double-dosage' of these genes. Many of these genes have highly similar but non-exact homologs on the Y-chromosome (XY gene pairs) to 'balance' gene expression between males and females. In particular, a growing body of work demonstrates the Y-chromosome has an active contribution to disease processes [40, 176, 177]. It was previously thought the expression of Y-chromosome genes is limited to the gonads; however, studies show that its expression can be elevated in non-reproductive tissues [178]. Furthermore, XY gene pairs include whole genome regulatory factors such as transcription factors (ZFX and ZFY), ribosomal subunits (RPS4X and RPS4Y1) and epigenetic histone demethylases (UTX and UTY) [1]. As such, these genes regulate the normal cellular and molecular processes at a whole genome level. As the sequences can differ up to 16% at

an amino acid level [137], this may lead to an imbalance or functional variations between the X and Y chromosome-linked versions. Earlier studies have shown that these genes regulate disease processes [45, 94, 179]. However, limited studies have attempted to explore and compare the functions of these genes to identify important regulatory mechanisms due to variations in the sequence of XY gene pairs and subsequent divergence in function. As females do not express the Y-linked version, this may lead to sex biased gene expression driving sexual dimorphism in disease susceptibility, severity and pathological mechanisms in asthma and COPD.

This notion informed the general aim of this thesis and project. To investigate the functions of candidate X and Y chromosome gene pairs (ZFX, ZFY, RPS4X, RPS4Y1, UTX and UTY) for their contribution in regulating hallmark features of asthma and COPD – inflammation, fibrosis and cell death. We further aimed to compare the functions of these genes to determine whether the Y chromosome-linked homolog can account for the double-dosage of the X chromosome-linked gene.

We generated CRISPR-Cas9 knockouts (KO) for the candidate XY gene pairs with multi-omics analyses to explore this aim. This methodology highlighted changes at a transcriptomic and protein level allowing the identification of novel functions and pathways for each gene. As a result, we have generated a valuable body of work that provides the impetus for future studies to further expand on the imbalance between X and Y chromosome gene pairs and their contribution to disease outcomes. The following three sections will outline our novel findings in inflammation, fibrosis and cell death associated with the candidate XY gene pairs.

8.1 XY gene pair regulation of the inflammatory response

Inflammation is a critical pathological feature in asthma and COPD, with TNFα established as a prominent elevated inflammatory stimulus [101, 180, 181]. As such, we have used TNFα as a proinflammatory stimulus to explore how the candidate genes regulate the immune response. CXCL8 and IL6 are major cytokines that activate the infiltration and activation of immune cells in asthma and COPD [182, 183], which have distinct regulatory pathways. Therefore, the observed suppression of CXCL8 and increased IL6 production in ZFX **(Chapter 5)** and RPS4Y1 **(Chapter 6)** knockout cell lines indicate that these genes function at the nexus of both regulatory pathways.

In ZFX knockouts, gene expression correlated with the relative change in protein production **(Chapter 5)**. In RPS4Y1 KOs, *CXCL8* gene expression was reduced, and *IL6* gene expression was similar to levels observed in wildtype cells **(Chapter 6)**. Therefore, in RPS4Y1 KOs, the translation of IL6 protein was explicitly increased. This data indicates that RPS4Y1 containing ribosomes can precisely regulate the translation of certain proteins [81]. In ZFX KO cells, both

RELB and NKRF transcription factor pathways were activated, which has been linked with the suppression of *CXCL8* gene expression [57] **(Chapter 5)**. Similarly, in RPS4Y1 KOs, NKRF protein is also increased. This, combined with suppression of the FOXA1 transcription factor pathway, which, when activated, increases CXCL8 mRNA production [109] **(Chapter 6)**. Thus, CXCL8 production is reduced. We also posit that ZFX modulates IL6 production through the CEBPA/CEBPB transcription factor pathway. CEBPA activity decreases in ZFX KOs while CEBPB increases **(Chapter 5)**. Overexpression of CEBPA attenuates IL6 production [58], whereas CEPBP functions synergistically with NF-κB (a key mediator of inflammatory responses [184]) to promote transcription of *IL6* [59].

When discussing the regulation of inflammation in both asthma and COPD, it is necessary to consider the viral inflammatory response. Common viruses such as human rhinovirus (HRV), respiratory syncytial virus (RSV) or influenza are found to be major drivers of exacerbations in both COPD and asthma [185]. Females with COPD or asthma experience more exacerbations and also demonstrate worse outcomes compared to males [186, 187]. Therefore, although beyond the scope of this thesis, it is pertinent to reflect on the roles these candidate gene pairs may play in modulating the inflammatory response in response to viral infections between the sexes. ZFX and RPS4Y1 were shown to strongly contribute to both CXCL8 and IL6 regulation **(Chapters 5 and 6)**. Both IL6 and CXCL8 are vital in the early response to viral infection [182, 188], with a recent study highlighting their increased gene expression after respiratory epithelial cells were infected with both HRV and RSV strains [189]. Further, ZFX has been associated with T-cell renewal and proliferation [44, 45], a vital immune cell involved in the viral inflammatory response [190]. The nasal epithelium, compared to the bronchial epithelium, demonstrates distinct responses to pathological stimuli [191]. In particular, nasal epithelium-derived cells present a more robust antiviral response than bronchial epithelial cells [192]. We observe a strong differential expression of *UTY* between the nasal and bronchial epithelium, which is affected by the presence of COPD **(Chapter 7)**. However, we observed a limited effect of knocking out UTY or UTX on the inflammatory response in our cellular models. This may be due to cell specificity as multiple studies have identified a role for UTX in modulating the effectiveness of virus-specific CD8+ T-cells [193], T-helper cells [194], or natural killer (NK) cells [195]. However, these studies do not compare the roles of UTX and UTY or explore whether the functions of these proteins differ depending on the location in the respiratory tract. As such, ZFX/ZFY and UTX/UTY present as sexually dimorphic genes that may contribute to sex differences observed in virus-induced exacerbations in asthma or COPD. Future studies need to investigate these candidate genes in the context of primary viral response cells, such as T-cells and NK cells, to potentially uncover new targets to mitigate viral infections in asthma and COPD.

These results indicate that ZFX and RPS4Y1 modulate complex signalling pathways and mechanisms contributing to the immune response. No effect on the immune response was reported for UTX and UTY **(Chapter 7)**. Future studies must holistically reveal how ZFX and RPS4Y1 mediate these transcription factor pathways. This notion is fundamental, as ZFY KOs did not indicate any change in the regulation of the immune response, identifying that ZFY function is distinct from ZFX in an inflammatory context. We also show that ZFX expression is greater in females than in males **(Chapter 5)**. Our data indicate that ZFX allows a more robust immune response, whilst there is a negligible contribution by ZFY. As such, this imbalance may contribute to a sexually dimorphic inflammatory response in asthma and COPD.

8.2 XY gene pair regulation of remodelling and fibrotic processes

Fibrosis is a complex and multifaceted process that manifests in asthma and COPD [52, 196], with poorly understood mechanisms driving this disease process. Fibrosis contributes to airway remodelling, ultimately causing a reduced breathing ability in patients via airway obstruction. The key features driving structural changes in the airways are a thickening of the epithelial layer through increased cell proliferation or cell migration, cell attachment, and altered extracellular matrix (ECM) deposition [196, 197]. Dysregulation of these features causes reduced breathing ability in patients due to a narrowing of the airway by thickened epithelial layers or reduced elasticity of the lungs [198]. The standard clinical method for assessing the physiological effect and disease progression uses FEV_1 and FVC lung function scores. Sex differences in these processes have been identified, where males indicate increased rates of fixed airflow obstruction [32], whilst another study reports greater fibrotic and remodelling changes in females [34]. As such, fibrosis and remodelling presents a key, clinically relevant process.

To assess regulation of migration and proliferation, we used the wound healing assay [199]. ZFX and RPS4Y1 KOs demonstrated a faster wound closure rate than wildtype cells **(Chapter 5 & 6)**. In contrast, ZFY KOs failed to close the wound after 72 hours, reinforcing that functional differences between ZFX and ZFY exist **(Chapter 5)**. ZFX knockouts also indicated a reduced proliferation rate with increased attachment to fibronectin, whilst only a trend towards decreased proliferation was seen for ZFY. Therefore, we postulated that knock out of ZFX increases cell migration, which was confirmed by pathway analysis. A slower proliferation rate leads to an inability of ZFY KOs to close the wound due to a breakdown in mitotic structures, which was also identified by pathway analysis. UTX and UTY single knockout cell lines indicate a trend towards a slower proliferation rate with doubling times of 24 and 25 hours, respectively, compared to wildtype cells with 23 hours. Double-knockout cells demonstrate a dramatically reduced proliferation rate, with a doubling time of 33 hours **(Chapter 7)**. This

highlights the ability of UTX and UTY to partially compensate for the loss of the other, indicating some overlapping functions between this XY gene pair.

Tenascin-C (TNC), a prominent ECM protein in asthma[200], was increased at a protein level in both ZFX and ZFY knockout cells **(Chapter 5)**. TNC is well known to promote cell migration and support cellular attachment [66]. This, combined with increased production of integrin subunit α4 (ITGA4) and reduction of integrin subunit β8 (ITGB8), further supports increased levels of cellular attachment and migration[201]. Notably, TNC usually is minimally expressed in adult tissues [65]. The stark increase of TNC from ZFX and RPS4Y1 KO cells indicates a novel function of these genes in regulating this ECM protein. *TNC* gene expression in RPS4Y1 KOs is unchanged compared to wildtype cells, whilst an increase in protein production is observed, whilst collagen α4 and fibronectin expression reflect protein production **(Chapter 6)**. This aligns with our previous observation for IL6, confirming that RPS4Y1 containing ribosomes preferentially promote or suppress the translation of specific mRNAs.

8.3 XY gene pair regulation of cell death in response to cigarette smoke extract

The regulation of cell death is an essential mechanism in maintaining homeostasis, where damaged or malfunctioning cells are eliminated [117]. Cigarette smoking is a pathological factor for asthma and COPD, causing worsened disease outcomes and severity [136, 202]. Dysregulation of cell death mechanisms persists in COPD even after smoking cessation [203]. Further, cells from asthma patients indicate increased sensitivity to cigarette smoke [204]. Sex differences in response to cigarette smoke are widely reported at the physiological and transcriptomic levels [28, 205, 206]. Therefore, the candidate XY gene pairs may contribute to sex differences in response to cigarette smoke and cell death mechanisms.

ZFY and UTY knockout cell lines demonstrated increased survival rates compared to wildtype cells in response to CSE exposure **(Chapters 5 & 7)**. RPS4Y1 KOs only indicated a trend towards reduced cell death **(Chapter 6)**. This data confirms that the Y chromosome is closely linked with the response to cigarette smoke [156]. Studies highlight that loss of the Y-chromosome promotes aberrant cell survival in males and is linked with an increased risk of disease and mortality [157, 207]. As identified by pathway analysis, we posit that increased ZFY KO survival is caused by a malformation of mitogenic structures such as the centrosomes **(Chapter 5)**. However, more detailed future studies are required to identify the precise mechanism by which ZFY regulates cell death processes.

Across all UTX and UTY knockout cells, we observed an apparent increase in expression of the gene for heat shock protein 90 alpha, class B, member 1 (*HSP90ab1)*, which is shown to prevent hyperoxic cell death [160] and be associated with COPD [158] **(Chapter 7)**. However,

this gene was also observed in UTX KOs, indicating that it cannot solely contribute to aberrant cell survival. We also identified that *CHORDC1* gene expression is explicitly increased in UTY and DKO cell lines, with the level of expression correlating with the amount of cell survival. CHORDC1 is a known co-chaperone factor that interacts with *HSP90ab1* [208]. Little is known about this interaction; thus, more work is required to uncover how these two factors interact with UTX and UTY to promote aberrant cell survival. Pathway analysis revealed the mitochondria are negatively enriched in UTY KO cells, highlighting a connection with *HSP90ab1*'s mechanism of action[163]. Importantly, we highlight through our investigation of primary cells that the mitochondria are dysregulated in COPD patient-derived cells. As such, we identify a mechanistic connection between UTY and cell death regulating factors in COPD **(Chapter 7)**. As UTY is a male-specific gene, this pathway may elucidate why sex differences in the smoking response and COPD outcomes exist.

8.4 Graphical summaries of the novel findings presented in this thesis

The following figures summarise the major findings of Chapters 5, 6 and 7.

8.4.1 ZFX and ZFY differentially regulate hallmark features of asthma – inflammation, fibrosis and death
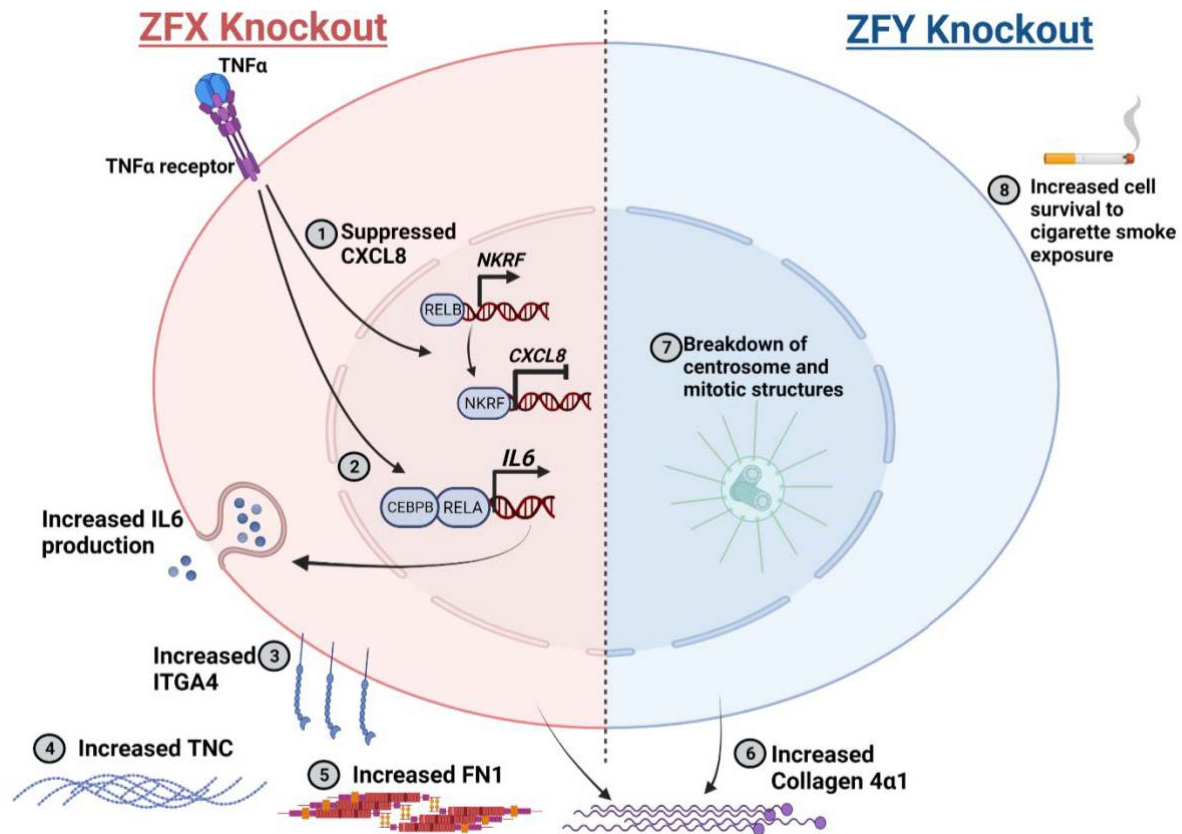
Figure 8.1: Chapter 5 summary - Altered cell processes in ZFX (1 – 6) and ZFY (6 – 8) knockout cells. *Created with BioRender.com*

1. Suppressed TNFα-induced expression of *CXCL8* by increased production of RELB and NKRF
2. Increased TNFα-induced expression of *IL6* by complexing of CEBPB and RELA
3. Increased expression and production of ITGA4
4. Increased expression and production of tenascin-C
5. Increased expression and production of fibronectin
6. Increased expression and production of collagen 4α1 (also in ZFY knockout cells)
7. Breakdown proteins needed for centrosome and microtubule recruitment
8. Increased cell survival in response to noxious stimuli in cigarette smoke extract

8.4.2 Novel regulatory role of RPS4Y1 in inflammation and fibrotic processes
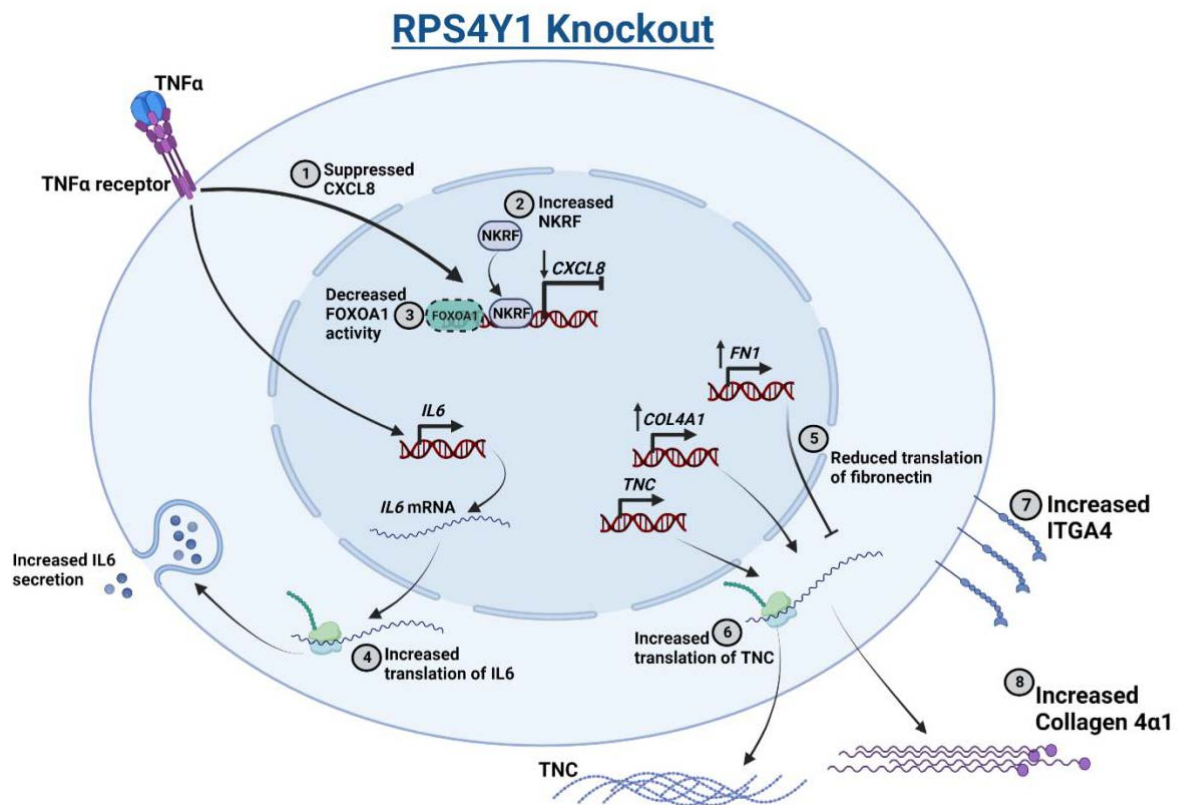
Figure 8.2: Chapter 6 summary - Gene and protein expression regulation in RPS4Y1 knockout cells. *Created with BioRender.com*

1. Suppressed TNFα-induced expression of *CXCL8* by increased production of NKRF protein (2) and reduced activity of FOXOA1 (3)
4. Increased TNFα-induced expression of *IL6* by increased translation of *IL6* mRNA into protein by RPS4Y1 negative ribosomes
5. Increased expression of *FN1* but reduced translation into protein by RPS4Y1 negative ribosomes
6. Normal expression of *TNC* with an increased translation of protein by RPS4Y1 negative ribosomes
7. Increased expression of ITGA4
8. Increased expression and protein production of collagen 4α1

8.4.3 The histone demethylase UTY uniquely contributes to COPD severity
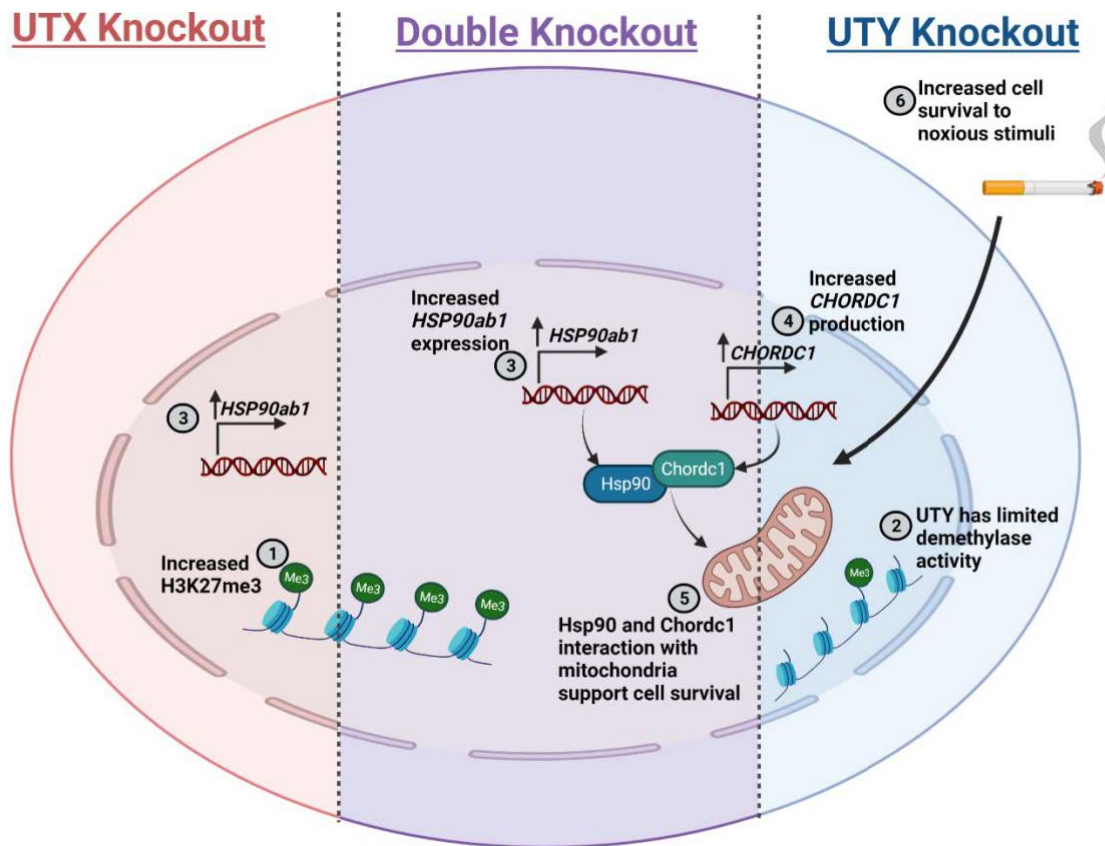
Figure 8.3: Chapter 7 summary - Cellular and molecular changes in UTX, UTY and double-knockout cells. *Created with BioRender.com*

1. Increase in H3K27me3 marks due to loss of demethylase activity (UTX KO and DKO cells)
2. Minimal change in H3K27me3 levels in UTY KOs
3. Increased expression of *HSP90ab1* (across all knockout cell lines)
4. Increase expression of *CHORDC1* (slight increase in UTY KO and greater increase in DKO cells)
5. HSP90 and Chordc1 interact with mitochondria to promote cell survival (UTY KO and DKO)
6. Slight increase in cell survival in response to noxious stimuli from UTY KO and greater increase from DKO cells

8.5 Future directions

This project provides substantial evidence and insight into the complex function of X and Y chromosome gene pairs. We have applied these findings in asthma and COPD, identifying pathologically relevant pathways that may contribute to sexual dimorphism in disease susceptibility, progression and prognosis. Nonetheless, significantly more studies are required to explore and confirm the mechanisms identified in this project.

The knockout models used throughout this thesis were generated in A549 cells, an alveolar basal epithelial cell type. As such, we developed a phenotype profile specific to this cell type. Epithelial cells are pivotal for multiple diseases contributing to disease development and progression [209]. Thus, we determined that epithelial cells function as a good model for understanding the fundamental functions of these genes. However, it is known, and we confirm that the candidate XY gene pairs demonstrate tissue and cell-specific expression patterns [2, 210]. Therefore, we believe conducting similar investigations in other cell types will be invaluable in holistically defining the function of these genes. Previous studies of these genes seemingly contradict some findings from the current body of work. Rashmi et al. found that UTY knockdown increased the expression of CXCL9 and CXCL10 cytokines in macrophages [211], which was not seen in our study.

Similarly, another study showed that IL6 production was reduced by RPS4Y1 knockdown in endothelial cells, whilst IL6 was increased in our investigation [103]. Asthma and COPD are diseases orchestrated by dysfunction of multiple cell types, from airway smooth muscle cells and fibroblasts to lymphocytes and macrophages [212], which contribute to sex differences. Therefore, future studies must establish how the XY gene pairs affect disease mechanisms.

The small interfering RNA (siRNA) knockdown model in primary cells isolated from disease-diagnosed patients with functional recovery studies will be incredibly valuable. This methodology and the multi-omics analysis used in this project could confirm many of the proposed mechanisms we have presented. Signalling pathways are highly complex, and a nexus of multiple factors function together to stimulate the production of one or multiple genes. As a result, there is significant redundancy in this pathway [213]. The current study's design did not allow any resolution to identify how the candidate XY gene pairs may function across different pathways. Using siRNA, recovery assays, and specific pathway inhibitors will allow for a more refined detection of mechanisms.

Furthermore, UTX and UTY are histone demethylases yet are known to have demethylase-independent functions [141] and can be associated with protein complexes to alter gene transcription [138]. However, few studies have explored this exciting and intricate mechanism. Protein tagging is a useful technique where changes in protein interactions can be tracked

and analysed. This study will provide valuable insight into the non-demethylase functions of UTX and UTY. Across all studies completed in **Chapters 4, 5 and 6,** we have generated an extensive list of differentially expressed genes. These studies highlight the crucial pathways mediated by XY gene pairs and provide an impetus for future studies of sexually dimorphic genetic and molecular factors. However, the suggested experiments discussed above will give concrete evidence for the precise mechanism of action and identify target pathways for developing novel clinical interventions.

The sex hormones, estrogen and testosterone, contribute to sex differences between males and females. Naturally, they present a critical factor in asthma, where disease susceptibility between males and females switches at puberty, but there is a lack of consensus regarding the effects of estrogen and testosterone in both sexes with asthma [214-216]. Further, both estrogen and testosterone are implicated in the progression and severity of multiple respiratory diseases [217-220]. Despite evidence that sex hormones modulate disease symptoms and outcomes, whether they contribute to pathogenesis and increased susceptibility remains unclear. Exploring this concept remained beyond the scope of the current project. However, UTX and UTY are described to interact with the estrogen receptor affecting gene transcription [221].

Further, RPS4Y1 is reported as a corticosteroid resistance gene [93]. Corticosteroids are an important class of medications in treating asthma that has been shown to be associated with estrogen [222, 223]. Therefore, sex chromosome-linked genes potentially present as a pivotal connection and factor in the function of sex hormones and sex differences in disease.

Using in vivo knockdown mouse models for the XY gene pairs will holistically examine the role of these genes at an organism level. As these genes are critical in embryonic development [78, 138, 224], using a global knockout model may result in lethality and invalid data. Using antisense oligonucleotide to induce gene knockdown, specifically in the lung, will present a model to analyse how these genes may contribute to the development of asthma and COPD. Furthermore, this model can be expanded to utilise the robust 'four core genome' model [225] to examine the contribution of sex hormones. These models, paired with *in vitro* models, will offer insights that begin to unravel the complex interaction between biological sex and disease.

8.6 Conclusions

This PhD thesis has identified novel contributions of ZFX, ZFY, RPS4Y1, UTX and UTY to the hallmark features of asthma and COPD – inflammation, fibrosis and cell death. Importantly, the function of these highly similar X and Y chromosome-linked gene pairs is not equivalent. As these genes modulate disease-relevant processes, divergent functions may lead to sex differences in the susceptibility and severity of disease in males and females. The data presented substantiate previous findings and contribute to the growing literature exploring sexually dimorphic processes. There is significant scope for future studies to continue elucidating and uncovering the factors contributing to sexual dimorphism in asthma and COPD.

Nonetheless, the current body of work provides a valuable base for future studies to expand upon. At the inception of this PhD project, the contribution of sex chromosome genes and consequences of sex-biased unequal gene expression was significantly understudied. In the following years, studies have dramatically increased the integration of biological sex as a contributing factor to disease outcomes. Cumulatively, the results and findings of this thesis provide novel insights into the role of genes encoded on the sex chromosomes in regulating critical pathological features of asthma and COPD. This vital data provides a foundation to identify new target pathways for the development of new, more effective treatments to improve patient outcomes.

Appendix A – Chapter 4 online supplementary data

Current-Smoking alters Gene Expression and DNA

Methylation in the Nasal Epithelium of Asthmatics

ONLINE DATA SUPPLEMENT

Supplementary File

Methods

Study Design

For this investigation, participants were enrolled as part of the 'Effects Of Extra-fine ParticLe HFA-Beclomethasone Versus Coarse Particle Treatment In Smokers and Ex-smokers with Asthma' (OLiVIA study). From nasal brushings taken from patients, differential gene expression and differential methylation analyses were conducted on samples which passed sample quality check procedures. From these analyses expression quantitative trait methylations (eQTM) analysis was completed. Separate analyses were conducted on the eQTM results investigating transcription factor binding proximity to methylation sites as well as longitudinal changes in methylation status. GSEA and GSVA analyses were used to compare our differential gene expression results to independent publically available gene expression datasets (Figure E1).

Sample Collection

Nasal brushings were taken by brushing the inferior turbinate of the nose. At the time of taking the nasal brush, patients did not use inhaled corticosteroids (ICS) or had stopped their ICS for 4-6 weeks. Nasal steroid use was permitted if constant dose was maintained throughout the study period.

Sample Processing

Nasal epithelial brushings were obtained using a cytology brush, sampling the inferior turbinate of the nose. RNA was isolated using a miRNeasy Mini Kit (Qiagen, Hilden, Germany). An Agilent 2100 BioAnalyser (Agilent, Santa Clara, CA, USA) was used to assess the quality of the RNA. Library preparation of mRNA was performed using a NEXTflex Rapid Directional RNA-Seq kit with poly (A) selection (Bio Scientific Corporation, Austin, TX, USA), using Caliper Sciclone NGS Workstation (PerkinElmer, Waltham, MA, USA). All samples were single-end sequenced (1x75bp) on Illumina HiSeq 2500 Sequencer.

RNA Sequencing

Alignment to the human genome build GRch38 was performed using STAR aligner (version 2.5.3a) and Ensemble gene annotation (release 88). HTSeq (version 0.6.1p1) was used for quantification. Quality control of gene expression was performed using R (version 3.4.3) with concordance between reported sex and sex-associated genes (*XIST* and Y-chromosome genes) confirmed. Genes with a mean expression ≤ 1 fragment per million across all samples were removed. The influence of technical variation was assessed via a principal component analysis (PCA) using the R package 'EDASeq' (version 2.12.0). The variation in library sizes was explored for gene expression among samples by making a relative log expression (RLE) plot. The library sizes cut-off ensured they would be larger than a quarter of the median of all library sizes. 26 samples were subsequently excluded due to low quality. Next, RNA-Seq data were log transformed and normalised using the 'voom' method from the R package 'limma' (version 3.34.0).

DNA Methylation Profiling

Genomic DNA was bisulfite-converted using the EZ DNA Methylation Gold Kit (#D5005, Zymo Research) and used for microarray-based DNA methylation analysis, performed at GenomeScan (GenomeScan B.V., The Netherlands) on the HumanMethylation850 BeadChip (Illumina, U.S.A). This array interrogates over 850,000 CpG sites representing about 99% of the RefSeq genes. The bisulfite-converted DNA was processed and hybridized to the HumanMethylation850 BeadChip (Illumina, U.S.A), according

E2

186

to the manufacturer's instructions. The BeadChip images were scanned on the iScan system and the data quality was assessed using the R script MethylAid (1) using default analysis settings.

Methylation Data Quality Assessment

Using the R packages 'minfi' (version 1.24.0), 'MethylAid' (version 1.12.0), and wateRmelon (version 1.22.1) methylation data quality was assessed. Raw intensity values read from IDAT files were converted to beta values using 'minfi'. 'MethylAid' identified bad quality samples, based on five diagnostic filter variables with the following thresholds: MU=10, OP=12, BS=11.50, HC=12.75, DP=0.95. MU - a median methylated and unmethylated log2 intensity smaller than [10.0]; Illustrated by the rotated MU plot; OP – an average log2 intensity of the expected signals in green and red channel of non-polymorphic controls smaller than [12.0]; illustrated by the overall sample-dependent control plot; BS – an average log2 intensity of converted Bisulfite Type 1 controls in green and red channel smaller than [11.50]; illustrated by the Bisulfite conversion control plot; HC – an average log2 intensity of High and Low hybridisation controls (green channel) smaller than [12.75]; illustrated by the overall sample-dependent control plot; DP – less than 95% of the probes above background signal [0.95]; illustrated by the detection p-value plot. Discordant samples were discarded as 'swapped' or 'contaminated' samples. Probes with high intensities indistinguishable from background, cross-reactive and type I probes with high intensity signals were discarded. All X- and Y-linked probes were discarded. Probes were filtered and raw beta-values were background corrected and normalised using the 'dasen' method from 'wateRmelon'.

Gene Expression Profiling and Mapping with 'in vitro' Datasets

Analysis of gene expression profiles in comparison with nasal gene expression generated in this study was completed using multiple publically available datasets from the Gene Expression Omnibus.

Gene set enrichment analysis (GSEA; version 3.0) was performed on the following two datasets:

GSE30660

This study isolated airway epithelial cells from the airways of patients, which were transformed into air-liquid interface (ALI) cultures in vitro. The ALI cultures were either challenged with cigarette smoke for 30min on four separate days or, exposed to no challenge. Differential gene expression analysis was completed between the two groups using a microarray; n=4.

GSE82137 (2)

Primary bronchial epithelial cells were isolated from a Caucasian 23-year old male non-smoker. Cells were transformed into ALI cultures in vitro and exposed to tobacco smoke and electronic cigarette smoke. Only the cigarette smoke data was compare to nasal epithelium gene expression as this exposure was relevant to our data. ALI cultures were exposed to whole tobacco smoke at 48 puffs in 48 min or air control, and rested for 24 hours before gene expression analysis was conducted using a microarray; n=3.

Gene set variation analysis (GSVA; version 1.36.0) was performed on the following datasets:

GSE113519

The adenocarcinoma human alveolar basal epithelial cells (A549) were treated with an Nrf2 siRNA to inhibit the NRF2 pathway. Differential gene expression analysis was completed by

E3

187

comparing siRNA treated cell gene expression to control RNA treated cell gene expression; n=3.

GSE109576 (3)

Comparison of gene expression in A549 cells exposed to AhR inhibitor CH223191 (10μM) or DMSO (0.01%) vehicle control for 6 hours; n=3.

GSE83364 (4)

A longitudinal analysis of nasal epithelium gene expression in response to smoking cessation. Total RNA was collected from the nasal epithelium of active smokers in smoking-cessation programs. Samples were collected at baseline, 4, 8, 16, and 14-weeks post smoking cessation; n=8. Patient demographics for the study population are included Table S2.

The following datasets were used for the mapping of DNA methylation sites with transcription factor binding sites:

GSE90550 (5)

Chromatin immunoprecipitation followed by next generation sequencing (ChIP-Seq) was completed on MCF-7 human breast cancer cells treated with vehicle (DMSO) or 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) for 24 hours; n=3.

GSE75812 (6)

ChIP-Seq analysis of human bronchial epithelium cells (BEAS-2B) treated with either sulforaphane (10uM) or DMSO (0.1% v/v) for 5 hours; n=2.

Longitudinal analysis of DNA methylation in bronchial biopsies after one-year smoking cessation

Bronchial biopsies were collected from 19 individuals (healthy= 7, COPD= 12) before and after 12 months smoking cessation followed by DNA extraction. DNA methylation was measured using Infinium-HumanMethylation850k sequencing and the differential methylation between before and after one year of smoking cessation was done using Bioconductor-limma package in R software version 3.5.1. A Benjamini–Hochberg corrected $p<0.05$ and the fold change $> |1.0|$ was considered as significant (7).

Cell Deconvolution Analysis

Single-cell RNA sequencing (scRNA-Seq) signatures for basal, ciliated and mucus-secretory cells (club and goblet cells) were utilized from our previously-published data to determine differences in cell-type composition, using mRNA expression levels (8). scRNA-Seq data from bronchial biopsies genes were selected which represented the unique profiles of each cell type, as previously explained (9). Due to similar expression profiles, club cell and goblet cell scRNA-Seq signatures were combined to generate a uniform scRNA-Seq signature of mucus-secretory cells. Bulk deconvolution was then conducted on our dataset, using the support vector regression (SVR) (10); the machine learning method implemented in CIBERSORT and, the non-negative least squares (NNLS) method (11). Resulting deconvolution predicted cell ratios were then compared within each dataset between smokers and non-smokers, using unpaired Mann-Whitney test.

E4

Results
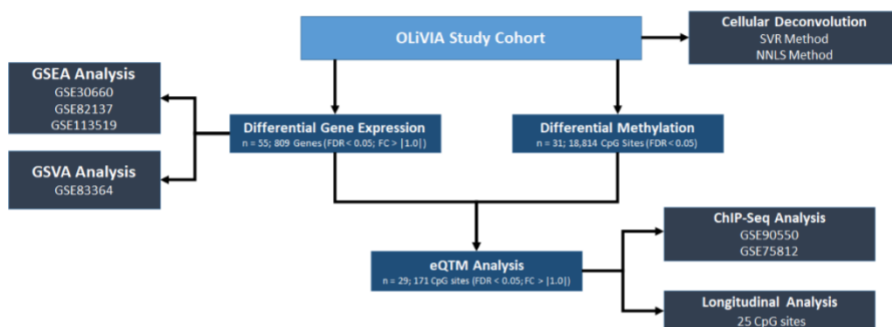
Figure E1: Schematic representation of the study design.
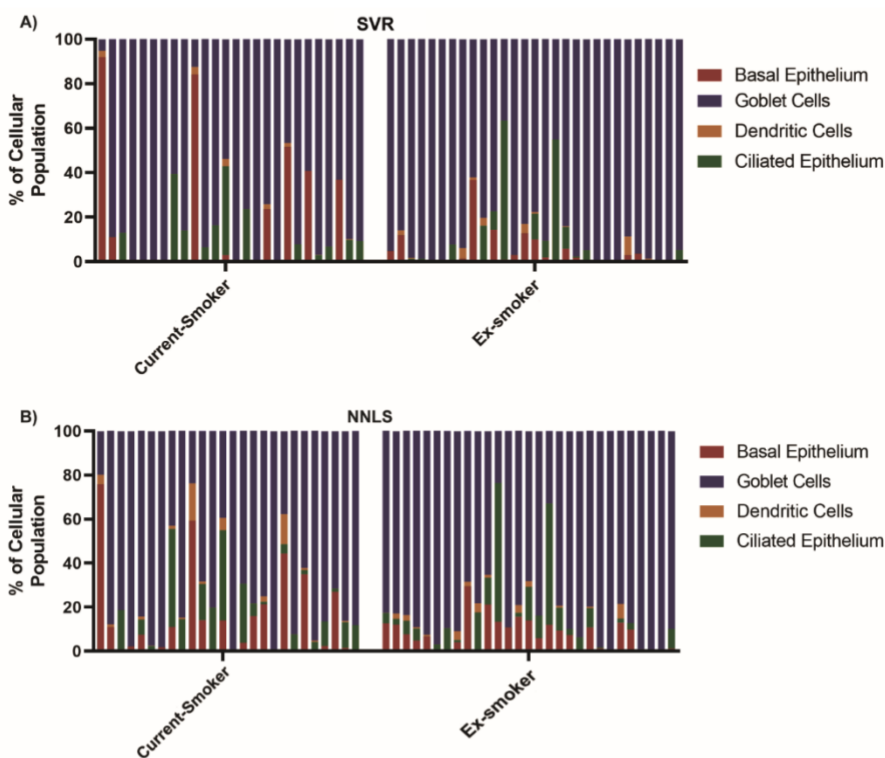
Figure E2: SVR and NNLS cell deconvolution analyses of cell compositions of patient nasal epithelium samples. Two methods were used to evaluate the cellular composition of patient samples (A) SVR and (B) NNLS. Goblet cells are the most common cell type with basal and ciliated epithelium and dendritic cells the other major cell types across the samples. No statistically significant differences were seen between current- and ex-smoker groups, by a Mann-Whitney test $p < 0.05$; n = 55.
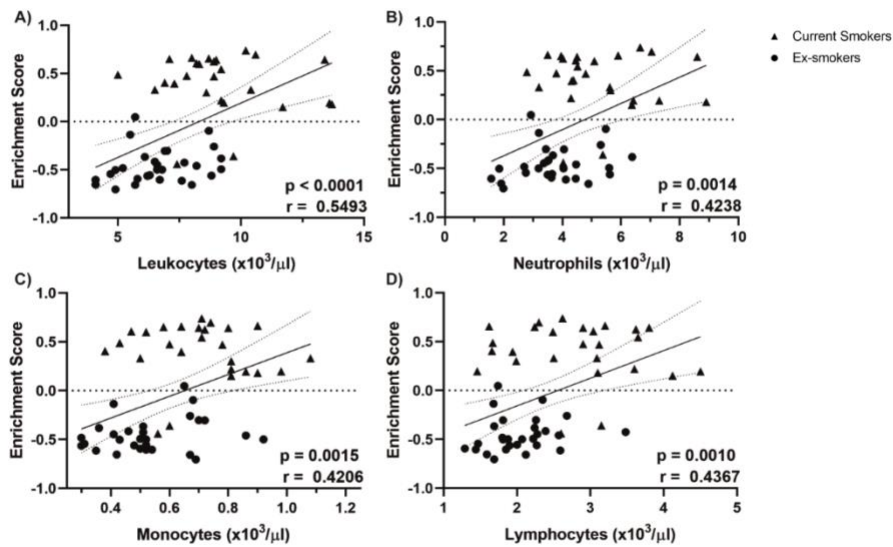
Figure E3: **Correlation between gene variation score of smoking related up-regulated genes and clinical parameters.** All figures indicate a gene enrichment score of zero by a dotted line, a solid regression line illustrating the direction of correlation between the X and Y variables, with 95% confidence intervals indicated by a dotted line. Data was analysed by non-parametric Spearman's correlation, r = correlation co-efficient. Statistical significance determined at $p < 0.05$; n = 26 (current smokers) and 29 (ex-smokers).
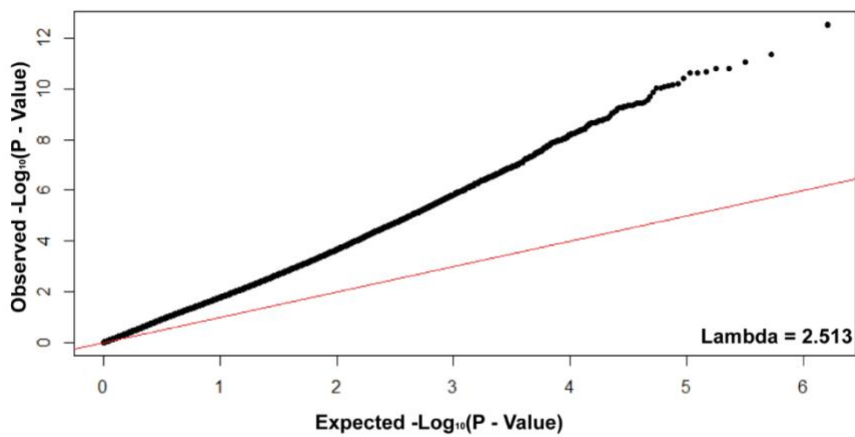
Figure E4: A quantile-quantile plot using nominal P-values of DNA methylation data from all patients.

Figure E5: Volcano Plot Visualisation of lower and higher methylation of CpG sites due to smoking by Log Fold Change. Volcano plot illustrating the CpG sites which are either lower (blue) and higher (red) methylated, $-\log_{10}$ (P-Value) plotted against the log (Fold Change). Vertical dotted lines indicate fold-change |1.0|. Genome-wide statistical significance is denoted in both figures by a solid black line at FDR adjusted p-value < 0.05; n = 14 (current smokers) and 17 (ex-smokers).

Figure E6: Mapping methylated CpG sites, transcription factor binding and transcript location. Diagrammatic representation of the location of significant smoking-related CpG sites with relevant transcription factors and transcripts of genes *CYP1B1* (A), *CYP1B1-AS1* (B) and *AHRR* (C).

E7

Figure E7: Frequency histogram of Boostrap analysis. Illustration of the number of CpG sites identified to be located in the gene regions investigated in the ChIP-Seq analysis. One-thousand permutations were completed using 171 randomly selected CpG sites from the total 797,128 CpG sites available on the array.

E8

Table E1: List of the 20 top significantly differentially expressed genes between ex-smokers and current smokers.

Gene Name	LogFC	p-value	FDR
CYP1A1	7.440636	2.89×10^{-23}	5.80×10^{-19}
CYP1B1	4.309097	8.15×10^{-18}	8.19×10^{-14}
GPX3	1.29748	5.86×10^{-14}	3.93×10^{-10}
ALDH3A1	1.098168	2.02×10^{-13}	1.02×10^{-09}
GLDN	1.435856	4.04×10^{-13}	1.62×10^{-09}
BPIFB2	3.18009	5.09×10^{-13}	1.70×10^{-09}
BPIFA2	1.523551	9.66×10^{-12}	2.49×10^{-08}
SAA1	-2.47157	9.93×10^{-12}	2.49×10^{-08}
SAA2	-2.49261	1.38×10^{-11}	3.09×10^{-08}
STATH	1.413974	9.11×10^{-11}	1.83×10^{-07}
CYP4X1	-1.03058	1.20×10^{-09}	2.19×10^{-06}
NQO1	1.06732	1.60×10^{-09}	2.68×10^{-06}
COL9A2	-1.01468	3.23×10^{-09}	5.00×10^{-06}
SLC7A11	1.316753	3.81×10^{-09}	5.47×10^{-06}
CYP1B1-AS1	2.312442	8.74×10^{-09}	1.16×10^{-05}
CHST4	-1.3559	2.30×10^{-08}	2.57×10^{-05}
MMP7	-2.33259	5.42×10^{-08}	5.58×10^{-05}
ABCA6	-2.09249	5.56×10^{-08}	5.58×10^{-05}
AHRR	1.945517	7.09×10^{-08}	6.79×10^{-05}
TIPARP	1.270451	7.73×10^{-08}	7.07×10^{-05}

LogFC = Log fold change; FDR = false discovery rate (Benjamini-Hochberg method)

Table E2: Summary demographic information for the in vivo smoking cessation model.

Characteristic	Male	Female
n	3	5
Age	58.67 (+/- 4.73)	39.20 (+/- 15.40)
Cigarettes per day (pre cessation)	12.50 (+/- 6.61)	14.20 (+/- 3.77)

Data is presented as mean +/- standard deviation. Information summarised from (4).

Table E3: List of the top 20 significantly differentially methylated CpG sites between ex-smokers and current smokers.

CpG Site	Delta Beta	LogFC	T-Value	FDR
cg12468966	0.14	1.01	15.24	2.42×10^{-07}
cg13754509	-0.02	-0.62	-13.33	1.78×10^{-06}
cg14837215	-0.08	-1.09	-12.92	2.20×10^{-06}
cg03470939	0.00	0.66	12.52	2.40×10^{-06}
cg16528039	-0.10	-1.08	-12.24	2.40×10^{-06}
cg24990564	-0.03	-0.59	-12.32	2.40×10^{-06}
cg07094194	0.12	0.56	11.67	4.80×10^{-06}
cg00303252	-0.01	1.14	12.26	2.40×10^{-06}
cg12651540	0.06	-1.06	-12.51	2.40×10^{-06}
cg27012574	0.02	-0.98	-11.42	4.80×10^{-06}
cg07017854	-0.02	0.65	11.53	4.80×10^{-06}
cg08448701	0.02	0.49	11.43	4.80×10^{-06}
cg26470501	0.01	-0.58	-11.45	4.80×10^{-06}
cg24165921	-0.06	-0.79	-11.59	4.80×10^{-06}
cg07211972	-0.04	0.64	11.21	6.50×10^{-06}
cg07440754	0.07	-0.50	-10.95	9.50×10^{-06}
cg12113316	-0.01	0.58	11.98	3.20×10^{-06}
cg19330536	0.02	0.57	10.65	1.32×10^{-05}
cg15787146	0.14	-0.62	-10.51	1.37×10^{-05}
cg11499323	-0.02	0.65	10.77	1.22×10^{-05}

LogFC = Log fold change; Delta Beta – difference in average beta methylation value between current and ex-smokers for the CpG site; FDR = false discovery rate (Benjamini-Hochberg method).

E10

194

Table E4: List of Top cis-eQTM CpG Sites for Genes Identified in Figure 4.

CpG	Gene	Statistic	P-Value	FDR	Beta
cg19949948	**ALDH3A1**	**-5.26**	1.91×10^{-05}	5.74×10^{-04}	**-0.91**
cg27638168	ALDH3A1	-4.14	3.47×10^{-04}	4.48×10^{-03}	-0.98
cg12258471	ALDH3A1	-4.03	4.60×10^{-04}	4.48×10^{-03}	-0.48
cg05575921	**AHRR**	**-4.01**	4.88×10^{-04}	2.10×10^{-02}	**-1.88**
cg21435924	**BPIFA2**	**4.29**	2.33×10^{-04}	6.52×10^{-03}	**1.84**
cg06978145	BPIFA2	3.94	5.81×10^{-04}	7.21×10^{-03}	0.46
cg13127741	BPIFA2	-3.83	7.73×10^{-04}	7.21×10^{-03}	-1.08
cg20385913	**CYP1A1**	**-3.52**	1.67×10^{-03}	4.48×10^{-02}	**-6.47**
cg25235648	CYP1A1	3.35	2.57×10^{-03}	4.48×10^{-02}	3.61
cg02780330	CYP1A1	-3.19	3.84×10^{-03}	4.48×10^{-02}	-12.65
cg11735238	**CYP1B1**	**-5.04**	3.38×10^{-05}	1.15×10^{-03}	**-8.99**
cg12802310	CYP1B1	-4.76	6.89×10^{-05}	1.17×10^{-03}	-5.54
cg05062676	CYP1B1	-4.16	3.27×10^{-04}	3.70×10^{-03}	-2.88
cg12802310	**CYP1B1-AS1**	**-3.86**	7.14×10^{-04}	1.84×10^{-02}	**-2.31**
cg11735238	CYP1B1-AS1	-3.69	1.08×10^{-03}	1.84×10^{-02}	-3.54
cg02162897	CYP1B1-AS1	-3.35	2.59×10^{-03}	2.93×10^{-02}	-0.75
cg13779050	**SLC7A11**	**2.71**	1.20×10^{-02}	3.81×10^{-02}	**1.01**
cg16988093	SLC7A11	2.51	1.91×10^{-02}	3.81×10^{-02}	0.50
cg05906530	SLC7A11	2.20	3.74×10^{-02}	4.98×10^{-02}	0.62

The most significant site for each gene is bolded. FDR = false discovery rate (Benjamini-Hochberg).

References

1.   Van Iterson M, Tobi EW, Slieker RC, Den Hollander W, Luijk R, Slagboom PE, et al. MethylAid: visual and interactive quality control of large Illumina 450k datasets. Bioinformatics. 2014;30(23):3435-7.
2.   Moses E, Wang T, Corbett S, Jackson GR, Drizik E, Perdomo C, et al. Molecular impact of electronic cigarette aerosol exposure in human bronchial epithelium. Toxicological Sciences. 2017;155(1):248-57.
3.   Procházková J, Strapáčová S, Svržková L, Andrysík Z, Hýždalová M, Hrubá E, et al. Adaptive changes in global gene expression profile of lung carcinoma A549 cells acutely exposed to distinct types of AhR ligands. Toxicology letters. 2018;292:162-74.
4.   Hijazi K, Malyszko B, Steiling K, Xiao X, Liu G, Alekseyev YO, et al. Tobacco-Related alterations in airway gene expression are rapidly reversed within weeks following smoking-cessation. Scientific reports. 2019;9(1):1-10.
5.   Yang SY, Ahmed S, Satheesh SV, Matthews J. Genome-wide mapping and analysis of aryl hydrocarbon receptor (AHR)-and aryl hydrocarbon receptor repressor (AHRR)-binding sites in human breast cancer cells. Archives of toxicology. 2018;92(1):225-40.
6.   Wang X, Campbell MR, Lacher SE, Cho H-Y, Wan M, Crowl CL, et al. A polymorphic antioxidant response element links NRF2/sMAF binding to enhanced MAPT expression and reduced risk of Parkinsonian disorders. Cell reports. 2016;15(4):830-42.
7.   Faiz A, Rathnayake SN, Vermeulen C, Timens W, Kooistra W, Oliver B, et al. Longitudinal effects of smoking cessation on DNA methylation in bronchial biopsies of COPD and asymptomatic smokers. Eur Respiratory Soc; 2019.
8.   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. 2019.
9.   Aliee H, Theis F. AutoGeneS: Automatic gene selection using multi-objective optimization for RNA-seq deconvolution. bioRxiv. 2020.

E11

10.   Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nature biotechnology. 2019;37(7):773-82.

11.   Mullen KM, van Stokkum IH. nnls: The Lawson-Hanson algorithm for non-negative least squares (NNLS), 2012. URL http://CRAN R-project org/package= nnls R package version.1.

Chapter 9 – References

1.   Bellott DW, Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Cho T-J, Koutseva N, Zaghlul S, Graves T, Rock S. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature* 2014: 508(7497): 494.
2.   Grath S, Parsch J. Sex-biased gene expression. *Annual Review of Genetics* 2016: 50: 29-44.
3.   Bradbury NA. All Cells Have a Sex: Studies of Sex Chromosome Function at the Cellular Level. Principles of Gender-Specific Medicine. Elsevier, 2017; pp. 269-290.
4.   Yung JA, Fuseini H, Newcomb DC. Sex hormones, gender and asthma. *Annals of allergy, asthma & immunology: official publication of the American College of Allergy, Asthma, & Immunology* 2018: 120(5): 488.
5.   Pouwels SD, Wiersma VR, Fokkema IE, Berg M, Ten Hacken NH, Van Den Berge M, Heijink I, Faiz A. Acute cigarette smoke-induced eQTL affects formyl peptide receptor expression and lung function. *Respirology* 2021: 26(3): 233-240.
6.   Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics* 2012: 13: 1-11.
7.   Suarez-Arnedo A, Torres Figueroa F, Clavijo C, Arbeláez P, Cruz JC, Muñoz-Camargo C. An image J plugin for the high throughput image analysis of in vitro scratch wound healing assays. *PloS one* 2020: 15(7): e0232565.
8.   Humphries MJ. Cell adhesion assays. *Extracellular Matrix Protocols* 2000: 279-285.
9.   Chen L, Ge Q, Tjin G, Alkhouri H, Deng L, Brandsma C-A, Adcock I, Timens W, Postma D, Burgess JK. Effects of cigarette smoke extract on human airway smooth muscle cells in COPD. *European Respiratory Journal* 2014: 44(3): 634-646.
10.   Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, Szcześniak MW, Gaffney DJ, Elo LL, Zhang X. A survey of best practices for RNA-seq data analysis. *Genome biology* 2016: 17(1): 1-19.
11.   Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 1995: 57(1): 289-300.
12.   Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Favera RD, Califano A. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. In: BMC bioinformatics; 2006: Springer; 2006. p. 1-15.
13.   Alvarez MJ, Shen Y, Giorgi FM, Lachmann A, Ding BB, Ye BH, Califano A. Functional characterization of somatic mutations in cancer using network-based inference of protein activity. *Nature genetics* 2016: 48(8): 838-847.
14.   Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC bioinformatics* 2013: 14(1): 1-15.
15.   Vieira Braga FA, Kar G, Berg M, Carpaij OA, Polanski K, Simon LM, Brouwer S, Gomes T, Hesse L, Jiang J. A cellular census of human lungs identifies novel cell states in health and in asthma. 2019.
16.   Sikkema L, Strobl DC, Zappia L, Madissoon E, Markov NS, Zaragosi L-E, Ansari M, Arguel M-J, Apperloo L, Becavin C. An integrated cell atlas of the human lung in health and disease. *bioRxiv* 2022: 2022.2003. 2010.483747.
17.   Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M. Integrated analysis of multimodal single-cell data. *Cell* 2021: 184(13): 3573-3587. e3529.
18.   Vermeulen CJ, Xu C-J, Vonk JM, Ten Hacken NH, Timens W, Heijink IH, Nawijn MC, Boekhoudt J, van Oosterhout AJ, Affleck K. Differential DNA methylation in bronchial biopsies between persistent asthma and asthma in remission. *European Respiratory Journal* 2020: 55(2).
19.   Broekema M, Volbeda F, Timens W, Dijkstra A, Lee N, Lee J, Lodewijk M, Postma D, Hylkema M, Ten Hacken N. Airway eosinophilia in remission and progression of asthma: accumulation with a fast decline of FEV1. *Respiratory medicine* 2010: 104(9): 1254-1262.

20.   Cox CA, Boudewijn IM, Vroegop SJ, Schokker S, Lexmond AJ, Frijlink HW, Hagedoorn P, Vonk JM, Farenhorst MP, Ten Hacken NH. Extrafine compared to non-extrafine particle inhaled corticosteroids in smokers and ex-smokers with asthma. *Respiratory Medicine* 2017: 130: 35-42.

21.   Rappsilber J, Mann M, Ishihama Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature protocols* 2007: 2(8): 1896-1906.

22.   Shah AD, Goode RJ, Huang C, Powell DR, Schittenhelm RB. LFQ-analyst: an easy-to-use interactive web platform to analyze and visualize label-free proteomics data preprocessed with MaxQuant. *Journal of proteome research* 2019: 19(1): 204-211.

23.   Chen W, Peng Y, Ma X, Kong S, Tan S, Wei Y, Zhao Y, Zhang W, Wang Y, Yan L. Integrated multi-omics reveal epigenomic disturbance of assisted reproductive technologies in human offspring. *EBioMedicine* 2020: 61: 103076.

24.   Deegan DF, Engel N. Sexual dimorphism in the age of genomics: how, when, where. *Frontiers in cell and developmental biology* 2019: 7: 186.

25.   Werner RJ, Schultz BM, Huhn JM, Jelinek J, Madzo J, Engel N. Sex chromosomes drive gene expression and regulatory dimorphisms in mouse embryonic stem cells. *Biology of sex differences* 2017: 8(1): 28.

26.   Maan AA, Eales J, Akbarov A, Rowland J, Xu X, Jobling MA, Charchar FJ, Tomaszewski M. The Y chromosome: a blueprint for men's health? *European Journal of Human Genetics* 2017: 25(11): 1181.

27.   Silkaitis K, Lemos B. Sex-biased chromatin and regulatory cross-talk between sex chromosomes, autosomes, and mitochondria. *Biology of sex differences* 2014: 5(1): 1-14.

28.   Reddy KD, Oliver BG. Sex-specific effects of in utero and adult tobacco smoke exposure. *American Journal of Physiology-Lung Cellular and Molecular Physiology* 2021: 320(1): L63-L72.

29.   Ekpruke CD, Silveyra P. Sex differences in airway remodeling and inflammation: clinical and biological factors. *Frontiers in Allergy* 2022: 3.

30.   Jenkins CR, Boulet L-P, Lavoie KL, Raherison-Semjen C, Singh D. Personalized treatment of asthma: the importance of sex and gender differences. *The Journal of Allergy and Clinical Immunology: In Practice* 2022.

31.   Asthma GIf. Global Strategy for Asthma Managemnet and Prevention: Global Initiative for Asthma; 2022.

32.   Chowdhury NU, Guntur VP, Newcomb DC, Wechsler ME. Sex and gender in asthma. *European Respiratory Review* 2021: 30(162).

33.   Uppal S, Verma S, Dhot P. Normal values of CD4 and CD8 lymphocyte subsets in healthy Indian adults and the effects of sex, age, ethnicity, and smoking. *Cytometry Part B: Clinical Cytometry: The Journal of the International Society for Analytical Cytology* 2003: 52(1): 32-36.

34.   Rasmussen F, Taylor DR, Flannery EM, Cowan JO, Greene JM, Herbison GP, Sears MR. Risk factors for airway remodeling in asthma manifested by a low postbronchodilator FEV1/vital capacity ratio: a longitudinal population study from childhood to adulthood. *American Journal of Respiratory and Critical Care Medicine* 2002: 165(11): 1480-1488.

35.   Osman M. Therapeutic implications of sex differences in asthma and atopy. *Archives of disease in childhood* 2003: 88(7): 587-590.

36.   Hibbert M, Lannigan A, Raven J, Landau L, Phelan P. Gender differences in lung growth. *Pediatric pulmonology* 1995: 19(2): 129-134.

37.   Charchar FJ, Bloomer LD, Barnes TA, Cowley MJ, Nelson CP, Wang Y, Denniff M, Debiec R, Christofidou P, Nankervis S. Inheritance of coronary artery disease in men: an analysis of the role of the Y chromosome. *The Lancet* 2012: 379(9819): 915-922.

38.   Wilson MA. The Y chromosome and its impact on health and disease. *Human Molecular Genetics* 2021: 30(R2): R296-R300.

39.   Umar S, Cunningham CM, Itoh Y, Moazeni S, Vaillancourt M, Sarji S, Centala A, Arnold AP, Eghbali M. The Y chromosome plays a protective role in experimental hypoxic pulmonary hypertension. *American Journal of Respiratory and Critical Care Medicine* 2018: 197(7): 952-955.

40.   Arnold AP. Y chromosome's roles in sex differences in disease. *Proc Natl Acad Sci U S A* 2017: 114(15): 3787-3789.

41.   Grants J, Flanagan E, Yee A, Romaniuk PJ. Characterization of the DNA binding activity of the ZFY zinc finger domain. *Biochemistry* 2010: 49(4): 679-686.

42.   North M, Sargent C, O'Brien J, Taylor K, Wolfe J, Affara N, Ferguson-Smith M. Comparison of ZFY and ZFX gene structure and analysis of alternative 3' untranslated regions of ZFY. *Nucleic acids research* 1991: 19(10): 2579-2586.

43.   Zhu Z, Li K, Xu D, Liu Y, Tang H, Xie Q, Xie L, Liu J, Wang H, Gong Y. ZFX regulates glioma cell proliferation and survival in vitro and in vivo. *Journal of neuro-oncology* 2013: 112(1): 17-25.

44.   Smith-Raska MR, Arenzana TL, D'Cruz LM, Khodadadi-Jamayran A, Tsirigos A, Goldrath AW, Reizis B. The transcription factor Zfx regulates peripheral T cell self-renewal and proliferation. *Frontiers in Immunology* 2018: 9: 1482.

45.   Weisberg SP, Smith-Raska MR, Esquilin JM, Zhang J, Arenzana TL, Lau CM, Churchill M, Pan H, Klinakis A, Dixon JE. ZFX controls propagation and prevents differentiation of acute T-lymphoblastic and myeloid leukemia. *Cell reports* 2014: 6(3): 528-540.

46.   Lee S, Prokopenko D, Kelly RS, Lutz S, Lasky-Su JA, Cho MH, Laurie C, Celedón JC, Lange C, Weiss ST. Zinc finger protein 33B demonstrates sex-interaction with atopy-related markers in childhood asthma. *European Respiratory Journal* 2022.

47.   Wu AC, Himes BE, Lasky-Su J, Litonjua A, Peters SP, Lima J, Kubo M, Tamari M, Nakamura Y, Qiu W. Inhaled corticosteroid treatment modulates ZNF432 gene variant's effect on bronchodilator response in asthmatics. *Journal of allergy and clinical immunology* 2014: 133(3): 723-728. e723.

48.   Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, Helmer Q, Tillander A, Ullemar V, Van Dongen J. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nature genetics* 2017: 49(12): 1752-1757.

49.   Li K, Zhu Z-C, Liu Y-J, Liu J-W, Wang H-T, Xiong Z-Q, Shen X, Hu Z-L, Zheng J. ZFX knockdown inhibits growth and migration of non-small cell lung carcinoma cell line H1299. *International journal of clinical and experimental pathology* 2013: 6(11): 2460.

50.   Xu L, Wang L, Cheng M. Identification of Genes and Pathways Associated with Sex in Non-smoking Lung Cancer Population. *Gene* 2022: 146566.

51.   Feng X, Zhou S, Cai W, Guo J. The miR-93-3p/ZFP36L1/ZFX axis regulates keratinocyte proliferation and migration during skin wound healing. *Molecular Therapy-Nucleic Acids* 2021: 23: 450-463.

52.   King GG, James A, Harkness L, Wark PA. Pathophysiology of severe asthma: We've only just started. *Respirology* 2018: 23(3): 262-271.

53.   Mick E, Tsitsiklis A, Spottiswoode N, Caldera S, Serpa PH, Detweiler AM, Neff N, Pisco AO, Li LM, Retallack H. Upper airway gene expression reveals a more robust innate and adaptive immune response to SARS-CoV-2 in children compared with older adults. *Research Square* 2021.

54.   Wood CL, Lane LC, Cheetham T. Puberty: Normal physiology (brief overview). *Best practice & research Clinical endocrinology & metabolism* 2019: 33(3): 101265.

55.   Ni W, Perez AA, Schreiner S, Nicolet CM, Farnham PJ. Characterization of the ZFX family of transcription factors that bind downstream of the start site of CpG island promoters. *Nucleic acids research* 2020: 48(11): 5986-6000.

56.   Silvestri M, Bontempelli M, Giacomelli M, Malerba M, Rossi G, Di Stefano A, Rossi A, Ricciardolo FLM. High serum levels of tumour necrosis factor-α and interleukin-8 in severe asthma: markers of systemic inflammation? *Clinical & Experimental Allergy* 2006: 36(11): 1373-1381.

57.   Ho S-C, Wu S-M, Feng P-H, Liu W-T, Chen K-Y, Chuang H-C, Chan Y-F, Kuo L-W, Lee K-Y. Noncanonical NF-κB mediates the suppressive effect of neutrophil elastase on IL-8/CXCL8 by inducing NKRF in human airway smooth muscle. *Scientific reports* 2017: 7(1): 1-11.

58.   Zhou J, Li H, Xia X, Herrera A, Pollock N, Reebye V, Sodergren MH, Dorman S, Littman BH, Doogan D. Anti-inflammatory Activity of MTL-CEBPA, a small activating RNA drug, in LPS-stimulated monocytes and humanized mice. *Molecular Therapy* 2019: 27(5): 999-1016.

59.   Matsusaka T, Fujikawa K, Nishio Y, Mukaida N, Matsushima K, Kishimoto T, Akira S. Transcription factors NF-IL6 and NF-kappa B synergistically activate transcription of the inflammatory cytokines, interleukin 6 and interleukin 8. *Proceedings of the National Academy of Sciences* 1993: 90(21): 10193-10197.

60.   Clifford RL, Patel JK, John AE, Tatler AL, Mazengarb L, Brightling CE, Knox AJ. CXCL8 histone H3 acetylation is dysfunctional in airway smooth muscle in asthma: regulation by BET. *American Journal of Physiology-Lung Cellular and Molecular Physiology* 2015: 308(9): L962-L972.

61.   Neveu WA, Allard JL, Raymond DM, Bourassa LM, Burns SM, Bunn JY, Irvin CG, Kaminsky DA, Rincon M. Elevation of IL-6 in the allergic asthmatic airway is independent of inflammation but associates with loss of central airway function. *Respiratory research* 2010: 11(1): 1-10.

62.   Khalaf H, Jass J, Olsson P-E. Differential cytokine regulation by NF-κB and AP-1 in Jurkat T-cells. *BMC immunology* 2010: 11(1): 1-12.

63.   Palmer MS, Berta P, Sinclair AH, Pym B, Goodfellow PN. Comparison of human ZFY and ZFX transcripts. *Proceedings of the National Academy of Sciences* 1990: 87(5): 1681-1685.

64.   Fang Q, Fu W-h, Yang J, Li X, Zhou Z-s, Chen Z-w, Pan J-h. Knockdown of ZFX suppresses renal carcinoma cell growth and induces apoptosis. *Cancer genetics* 2014: 207(10-12): 461-466.

65.   Chiovaro F, Chiquet-Ehrismann R, Chiquet M. Transcriptional regulation of tenascin genes. *Cell adhesion & migration* 2015: 9(1-2): 34-47.

66.   Chiquet-Ehrismann R, Tucker RP. Tenascins and the importance of adhesion modulation. *Cold Spring Harbor perspectives in biology* 2011: 3(5): a004960.

67.   Chiquet-Ehrismann R, Kalla P, Pearson CA, Beck K, Chiquet M. Tenascin interferes with fibronectin action. *Cell* 1988: 53(3): 383-390.

68.   Mills JT, Schwenzer A, Marsh EK, Edwards MR, Sabroe I, Midwood KS, Parker LC. Airway epithelial cells generate pro-inflammatory tenascin-C and small extracellular vesicles in response to TLR3 stimuli and rhinovirus infection. *Frontiers in immunology* 2019: 1987.

69.   Maqbool A, Spary EJ, Manfield IW, Ruhmann M, Zuliani-Alvarez L, Gamboa-Esteves FO, Porter KE, Drinkhill MJ, Midwood KS, Turner NA. Tenascin C upregulates interleukin-6 expression in human cardiac myofibroblasts via toll-like receptor 4. *World journal of cardiology* 2016: 8(5): 340.

70.   Banday AR, Papenberg BW, Prokunina-Olsson L. When the Smoke Clears m6A from a Y Chromosome–Linked lncRNA, Men Get an Increased Risk of Cancer. *Cancer Research* 2020: 80(13): 2718-2719.

71.   van den Berge M, Brandsma C-A, Faiz A, de Vries M, Rathnayake SN, Paré PD, Sin DD, Bossé Y, Laviolette M, Nickle DC. Differential lung tissue gene expression in males and females: implications for the susceptibility to develop COPD. *European Respiratory Journal* 2019: 54(1).

72.   Irimie AI, Braicu C, Cojocneanu R, Magdo L, Onaciu A, Ciocan C, Mehterov N, Dudea D, Buduru S, Berindan-Neagoe I. Differential effect of smoking on gene expression in head and neck cancer patients. *International journal of environmental research and public health* 2018: 15(7): 1558.

73.   Davoli T, Xu AW, Mengwasser KE, Sack LM, Yoon JC, Park PJ, Elledge SJ. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* 2013: 155(4): 948-962.

74.   Arenzana TL, Smith-Raska MR, Reizis B. Transcription factor Zfx controls BCR-induced proliferation and survival of B lymphocytes. *Blood, The Journal of the American Society of Hematology* 2009: 113(23): 5857-5867.

75.   Lafontaine DL, Tollervey D. The function and synthesis of ribosomes. *Nature Reviews Molecular Cell Biology* 2001: 2(7): 514-520.

76.   Dolezal JM, Dash AP, Prochownik EV. Diagnostic and prognostic implications of ribosomal protein transcript expression patterns in human cancers. *BMC cancer* 2018: 18(1): 1-14.

77.   Zhou X, Liao W-J, Liao J-M, Liao P, Lu H. Ribosomal proteins: functions beyond the ribosome. *Journal of molecular cell biology* 2015: 7(2): 92-104.

78.   Andrés O, Kellermann T, López-Giráldez F, Rozas J, Domingo-Roura X, Bosch M. RPS4Y gene family evolution in primates. *BMC evolutionary biology* 2008: 8(1): 1-12.

79.   Fisher EM, Beer-Romero P, Brown LG, Ridley A, McNeil JA, Lawrence JB, Willard HF, Bieber FR, Page DC. Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and possible implications for Turner syndrome. *Cell* 1990: 63(6): 1205-1218.

80.   Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 2003: 423(6942): 825.

81.   Lopes AM, Miguel RN, Sargent CA, Ellis PJ, Amorim A, Affara NA. The human RPS4 paralogue on Yq11. 223 encodes a structurally conserved ribosomal protein and is preferentially expressed during spermatogenesis. *BMC molecular biology* 2010: 11(1): 1-12.

82.   Bergen AW, Pratt M, Mehlman PT, Goldman D. Evolution of RPS4Y. *Molecular biology and evolution* 1998: 15(11): 1412-1419.

83.   Zinn AR, Alagappan RK, Brown LG, Wool I, Page DC. Structure and function of ribosomal protein S4 genes on the human and mouse sex chromosomes. *Molecular and cellular biology* 1994: 14(4): 2485-2492.

84.   Watanabe M, Zinn AR, Page DC, Nishimoto T. Functional equivalence of human X–and Y–encoded isoforms of ribosomal protein S4 consistent with a role in Turner syndrome. *Nature genetics* 1993: 4(3): 268-271.

85.   Johansson M, Ieong K-W, Trobro S, Strazewski P, Åqvist J, Pavlov MY, Ehrenberg M. pH-sensitivity of the ribosomal peptidyl transfer reaction dependent on the identity of the A-site aminoacyl-tRNA. *Proceedings of the National Academy of Sciences* 2011: 108(1): 79-84.

86.   Stephen C, Mishanina TV. Alkaline pH has an unexpected effect on transcriptional pausing during synthesis of the Escherichia coli pH-responsive riboswitch. *Journal of Biological Chemistry* 2022: 298(9).

87.   Fuseini H, Newcomb DC. Mechanisms driving gender differences in asthma. *Current allergy and asthma reports* 2017: 17(3): 1-9.

88.   Lau Y-FC. Y chromosome in health and diseases. *Cell & bioscience* 2020: 10(1): 1-10.

89.   Tsofack SP, Meunier L, Sanchez L, Madore J, Provencher D, Mes-Masson A-M, Lebel M. Low expression of the X-linked ribosomal protein S4 in human serous epithelial ovarian cancer is associated with a poor prognosis. *BMC cancer* 2013: 13(1): 1-12.

90.   Bi G, Zhu D, Bian Y, Huang Y, Zhan C, Yang Y, Wang Q. Knockdown of GTF2E2 inhibits the growth and progression of lung adenocarcinoma via RPS4X in vitro and in vivo. *Cancer cell international* 2021: 21(1): 1-13.

91.   Paquet ÉR, Hovington H, Brisson H, Lacombe C, Larue H, Têtu B, Lacombe L, Fradet Y, Lebel M. Low level of the X-linked ribosomal protein S4 in human urothelial carcinomas is associated with a poor prognosis. *Biomarkers in medicine* 2015: 9(3): 187-197.

92.   Zhou P, Xiang Cx, Wei Jf. The clinical significance of spondin 2 eccentric expression in peripheral blood mononuclear cells in bronchial asthma. *Journal of Clinical Laboratory Analysis* 2021: 35(6): e23764.

93.   Chang R, Chen L, Su G, Du L, Qin Y, Xu J, Tan H, Zhou C, Cao Q, Yuan G. Identification of Ribosomal Protein S4, Y-Linked 1 as a cyclosporin A plus corticosteroid resistance gene. *Journal of autoimmunity* 2020: 112: 102465.

94.   Hu Y, Hu Y, Xiao Y, Wen F, Zhang S, Liang D, Su L, Deng Y, Luo J, Ou J. Genetic landscape and autoimmunity of monocytes in developing Vogt–Koyanagi–Harada disease. *Proceedings of the National Academy of Sciences* 2020: 117(41): 25712-25721.

95.   Zhao H-c, Chen C-z, Song H-q, Wang X-x, Zhang L, Zhao H-l, He J-f. Single-cell RNA Sequencing Analysis Reveals New Immune Disorder Complexities in Hypersplenism. *Frontiers in immunology* 2022: 13: 921900.

96.   Chen G, Volmer AS, Wilkinson KJ, Deng Y, Jones LC, Yu D, Bustamante-Marin XM, Burns KA, Grubb BR, O'Neal WK. Role of Spdef in the regulation of Muc5b expression in the airways of naive and mucoobstructed mice. *American journal of respiratory cell and molecular biology* 2018: 59(3): 383-396.

97.   Qian B, Yao Z, Yang Y, Li N, Wang Q. Downregulation of SDCBP inhibits cell proliferation and induces apoptosis by regulating PI3K/AKT/mTOR pathway in gastric carcinoma. *Biotechnology and Applied Biochemistry* 2022: 69(1): 240-247.

98.   Ge Q, Zeng Q, Tjin G, Lau E, Black JL, Oliver BG, Burgess JK. Differential deposition of fibronectin by asthmatic bronchial epithelial cells. *American Journal of Physiology-Lung Cellular and Molecular Physiology* 2015: 309(10): L1093-L1102.

99.   Gremlich S, Roth-Kleiner M, Equey L, Fytianos K, Schittny JC, Cremona TP. Tenascin-C inactivation impacts lung structure and function beyond lung development. *Scientific reports* 2020: 10(1): 1-13.

100.   Liu L, Stephens B, Bergman M, May A, Chiang T. Role of collagen in airway mechanics. *Bioengineering* 2021: 8(1): 13.

101.   Yu H, Guo W, Liu Y, Wang Y. Immune Characteristics Analysis and Transcriptional Regulation Prediction Based on Gene Signatures of Chronic Obstructive Pulmonary Disease. *International journal of chronic obstructive pulmonary disease* 2021: 16: 3027.

102.   Liu C, Zhang X, Xiang Y, Qu X, Liu H, Liu C, Tan M, Jiang J, Qin X. Role of epithelial chemokines in the pathogenesis of airway inflammation in asthma. *Molecular medicine reports* 2018: 17(5): 6935-6941.

103.   Chen Y, Chen Y, Tang C, Zhao Q, Xu T, Kang Q, Jiang B, Zhang L. RPS4Y1 Promotes High Glucose-Induced Endothelial Cell Apoptosis and Inflammation by Activation of the p38 MAPK Signaling. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy* 2021: 14: 4523.

104.   Bonser LR, Erle DJ. Airway mucus and asthma: the role of MUC5AC and MUC5B. *Journal of clinical medicine* 2017: 6(12): 112.

105.   Chen G, Korfhagen TR, Karp CL, Impey S, Xu Y, Randell SH, Kitzmiller J, Maeda Y, Haitchi HM, Sridharan A. Foxa3 induces goblet cell metaplasia and inhibits innate antiviral immunity. *American journal of respiratory and critical care medicine* 2014: 189(3): 301-313.

106.   Rajavelu P, Chen G, Xu Y, Kitzmiller JA, Korfhagen TR, Whitsett JA. Airway epithelial SPDEF integrates goblet cell differentiation and pulmonary Th2 inflammation. *The Journal of clinical investigation* 2015: 125(5): 2021-2031.

107.   Park K-S, Korfhagen TR, Bruno MD, Kitzmiller JA, Wan H, Wert SE, Hershey GKK, Chen G, Whitsett JA. SPDEF regulates goblet cell hyperplasia in the airway epithelium. *The Journal of clinical investigation* 2007: 117(4): 978-988.

108.   McKay KO, Hogg JC. The contribution of airway structure to early childhood asthma. *Medical journal of Australia* 2002: 177: S45-S47.

109.   Fu X, Jeselsohn R, Pereira R, Hollingsworth EF, Creighton CJ, Li F, Shea M, Nardone A, De Angelis C, Heiser LM. FOXA1 overexpression mediates endocrine resistance by altering the ER transcriptome and IL-8 expression in ER-positive breast cancer. *Proceedings of the National Academy of Sciences* 2016: 113(43): E6600-E6609.

110.   Chen X, Tong C, Li H, Peng W, Li R, Luo X, Ge H, Ran Y, Li Q, Liu Y. Dysregulated expression of RPS4Y1 (ribosomal protein S4, Y-linked 1) impairs STAT3 (signal transducer and activator of transcription 3) signaling to suppress trophoblast cell migration and invasion in preeclampsia. *Hypertension* 2018: 71(3): 481-490.

111.   David S, Edwards CW. Forced expiratory volume. StatPearls [Internet]. StatPearls Publishing, 2021.

112.   Firoozi F, Lemière C, Beauchesne M-F, Forget A, Blais L. Development and validation of database indexes of asthma severity and control. *Thorax* 2007: 62(7): 581-587.

113.   Wu D-D, Song J, Bartel S, Krauss-Etschmann S, Rots MG, Hylkema MN. The potential for targeted rewriting of epigenetic marks in COPD as a new therapeutic approach. *Pharmacology & therapeutics* 2018: 182: 1-14.

114.   (GOLD) GIfCOLD. Pocket Guide to COPD Diagnosis, Management and Prevention: Global Intiative for Chronic Obstructive Lung Disease; 2019.

115.	Adeloye D, Song P, Zhu Y, Campbell H, Sheikh A, Rudan I, Unit NRGRH. Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis. *The Lancet Respiratory Medicine* 2022.

116.	Capistrano SJ, Van Reyk D, Chen H, Oliver BG. Evidence of biomass smoke exposure as a causative factor for the development of COPD. *Toxics* 2017: 5(4): 36.

117.	Demedts IK, Demoor T, Bracke KR, Joos GF, Brusselle GG. Role of apoptosis in the pathogenesis of COPD and pulmonary emphysema. *Respiratory research* 2006: 7(1): 1-10.

118.	Lu Z, Van Eeckhoutte HP, Liu G, Nair PM, Jones B, Gillis CM, Nalkurthi BC, Verhamme F, Buyle-Huybrecht T, Vandenabeele P. Necroptosis signaling promotes inflammation, airway remodeling, and emphysema in chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine* 2021: 204(6): 667-681.

119.	Giordano RJ, Lahdenranta J, Zhen L, Chukwueke U, Petrache I, Langley RR, Fidler IJ, Pasqualini R, Tuder RM, Arap W. Targeted Induction of Lung Endothelial Cell Apoptosis Causes Emphysema-like Changes in the Mouse∗. *Journal of Biological Chemistry* 2008: 283(43): 29447-29460.

120.	Hou H, Cheng S, Liu H, Yang F, Wang H, Yu C. Elastase induced lung epithelial cell apoptosis and emphysema through placenta growth factor. *Cell death & disease* 2013: 4(9): e793-e793.

121.	Laniado-Laborín R. Smoking and chronic obstructive pulmonary disease (COPD). Parallel epidemics of the 21st century. *International journal of environmental research and public health* 2009: 6(1): 209-224.

122.	Salvi SS, Barnes PJ. Chronic obstructive pulmonary disease in non-smokers. *The lancet* 2009: 374(9691): 733-743.

123.	Silverman EK, Weiss ST, Drazen JM, Chapman HA, Carey V, Campbell EJ, Denish P, Silverman RA, Celedon JC, Reilly JJ. Gender-related differences in severe, early-onset chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine* 2000: 162(6): 2152-2158.

124.	Barnes PJ. Sex differences in chronic obstructive pulmonary disease mechanisms. American Thoracic Society, 2016.

125.	Foreman MG, Zhang L, Murphy J, Hansel NN, Make B, Hokanson JE, Washko G, Regan EA, Crapo JD, Silverman EK. Early-onset chronic obstructive pulmonary disease is associated with female sex, maternal factors, and African American race in the COPDGene Study. *American journal of respiratory and critical care medicine* 2011: 184(4): 414-420.

126.	Eisner MD, Balmes J, Katz PP, Trupin L, Yelin EH, Blanc PD. Lifetime environmental tobacco smoke exposure and the risk of chronic obstructive pulmonary disease. *Environmental Health* 2005: 4(1): 7.

127.	Tam A, Churg A, Wright JL, Zhou S, Kirby M, Coxson HO, Lam S, Man SP, Sin DD. Sex differences in airway remodeling in a mouse model of chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine* 2016: 193(8): 825-834.

128.	Gan WQ, Man SP, Postma DS, Camp P, Sin DD. Female smokers beyond the perimenopausal period are at increased risk of chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Respiratory research* 2006: 7(1): 1-9.

129.	DeMeo DL. Sex and gender omic biomarkers in men and women with COPD: considerations for precision medicine. *Chest* 2021: 160(1): 104-113.

130.	Panning B. X-chromosome inactivation: the molecular basis of silencing. *Journal of biology* 2008: 7(8): 1-4.

131.	Martin C, Zhang Y. The diverse functions of histone lysine methylation. *Nature reviews Molecular cell biology* 2005: 6(11): 838-849.

132.	Day CA, Hinchcliffe EH, Robinson JP. H3K27me3 in Diffuse Midline Glioma and Epithelial Ovarian Cancer: Opposing Epigenetic Changes Leading to the Same Poor Outcomes. *Cells* 2022: 11(21): 3376.

133.   Alford SH, Toy K, Merajver SD, Kleer CG. Increased risk for distant metastasis in patients with familial early-stage breast cancer and high EZH2 expression. *Breast cancer research and treatment* 2012: 132(2): 429-437.

134.   Byrd AL, Qu X, Lukyanchuk A, Liu J, Chen F, Naughton KJ, DuCote TJ, Song X, Bowman HC, Zhao Y, Edgin AR, Wang C, Liu J, Brainson CF. Dysregulated polycomb repressive complex 2 contributes to chronic obstructive pulmonary disease by rewiring stem cell fate. *Stem Cell Reports* 2022.

135.   Anzalone G, Arcoleo G, Bucchieri F, Montalbano AM, Marchese R, Albano GD, Di Sano C, Moscato M, Gagliardo R, Ricciardolo FL. Cigarette smoke affects the onco-suppressor DAB2IP expression in bronchial epithelial cells of COPD patients. *Scientific reports* 2019: 9(1): 1-14.

136.   Sundar IK, Nevid MZ, Friedman AE, Rahman I. Cigarette smoke induces distinct histone modifications in lung cells: implications for the pathogenesis of COPD and lung cancer. *Journal of proteome research* 2014: 13(2): 982-996.

137.   Xu J, Deng X, Watkins R, Disteche CM. Sex-specific differences in expression of histone demethylases Utx and Uty in mouse brain and neurons. *Journal of Neuroscience* 2008: 28(17): 4521-4527.

138.   Shpargel KB, Sengoku T, Yokoyama S, Magnuson T. UTX and UTY demonstrate histone demethylase-independent function in mouse embryonic development. 2012.

139.   Lan F, Bayliss PE, Rinn JL, Whetstine JR, Wang JK, Chen S, Iwase S, Alpatov R, Issaeva I, Canaani E. A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* 2007: 449(7163): 689-694.

140.   Hong S, Cho Y-W, Yu L-R, Yu H, Veenstra TD, Ge K. Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proceedings of the National Academy of Sciences* 2007: 104(47): 18439-18444.

141.   Walport LJ, Hopkinson RJ, Vollmar M, Madden SK, Gileadi C, Oppermann U, Schofield CJ, Johansson C. Human UTY (KDM6C) is a male-specific Nϵ-methyl lysyl demethylase. *Journal of Biological Chemistry* 2014: 289(26): 18302-18313.

142.   Welstead GG, Creyghton MP, Bilodeau S, Cheng AW, Markoulaki S, Young RA, Jaenisch R. X-linked H3K27me3 demethylase Utx is required for embryonic development in a sex-specific manner. *Proceedings of the National Academy of Sciences* 2012: 109(32): 13004-13009.

143.   GOLD. Global Strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease 2023 report; 2023.

144.   Krimmer D, Ichimaru Y, Burgess J, Black J, Oliver B. Exposure to biomass smoke extract enhances fibronectin release from fibroblasts. *PLoS One* 2013: 8(12).

145.   Wang B, Chan Y-L, Li G, Ho KF, Anwer AG, Smith BJ, Guo H, Jalaludin B, Herbert C, Thomas PS. Maternal particulate matter exposure impairs lung health and is associated with mitochondrial damage. *Antioxidants* 2021: 10(7): 1029.

146.   Mori M, Hitora T, Nakamura O, Yamagami Y, Horie R, Nishimura H, Yamamoto T. Hsp90 inhibitor induces autophagy and apoptosis in osteosarcoma cells. *International journal of oncology* 2015: 46(1): 47-54.

147.   Gažová I, Lengeling A, Summers KM. Lysine demethylases KDM6A and UTY: The X and Y of histone demethylation. *Molecular genetics and metabolism* 2019: 127(1): 31-44.

148.   Gottwald EM, Duss M, Bugarski M, Haenni D, Schuh CD, Landau EM, Hall AM. The targeted anti-oxidant MitoQ causes mitochondrial swelling and depolarization in kidney tissue. *Physiological reports* 2018: 6(7): e13667.

149.   Gower AC, Steiling K, Brothers JF, Lenburg ME, Spira A. Transcriptomic studies of the airway field of injury associated with smoking-related lung disease. *Proceedings of the American Thoracic Society* 2011: 8(2): 173-179.

150.   Fazleen A, Wilkinson T. Early COPD: current evidence for diagnosis and management. *Therapeutic advances in respiratory disease* 2020: 14: 1753466620942128.

151.   Baraldo S, Turato G, Badin C, Bazzan E, Beghé B, Zuin R, Calabrese F, Casoni G, Maestrelli P, Papi A. Neutrophilic infiltration within the airway smooth muscle in patients with COPD. *Thorax* 2004: 59(4): 308-312.

152.   Cunningham CM, Li M, Ruffenach G, Doshi M, Aryan L, Hong J, Park J, Hrncir H, Medzikovic L, Umar S. Y-Chromosome Gene, Uty, Protects Against Pulmonary Hypertension by Reducing Proinflammatory Chemokines. *American Journal of Respiratory and Critical Care Medicine* 2022(ja).

153.   Wang C, Lee J-E, Cho Y-W, Xiao Y, Jin Q, Liu C, Ge K. UTX regulates mesoderm differentiation of embryonic stem cells independent of H3K27 demethylase activity. *Proceedings of the National Academy of Sciences* 2012: 109(38): 15324-15329.

154.   Tran N, Broun A, Ge K. Lysine demethylase KDM6A in differentiation, development, and cancer. *Molecular and cellular biology* 2020: 40(20): e00341-00320.

155.   Cáceres A, Jene A, Esko T, Pérez-Jurado LA, González JR. Extreme downregulation of chromosome Y and cancer risk in men. *JNCI: Journal of the National Cancer Institute* 2020: 112(9): 913-920.

156.   Dumanski JP, Rasi C, Lönn M, Davies H, Ingelsson M, Giedraitis V, Lannfelt L, Magnusson PK, Lindgren CM, Morris AP. Smoking is associated with mosaic loss of chromosome Y. *Science* 2015: 347(6217): 81-83.

157.   Noveski P, Madjunkova S, Sukarova Stefanovska E, Matevska Geshkovska N, Kuzmanovska M, Dimovski A, Plaseska-Karanfilska D. Loss of Y chromosome in peripheral blood of colorectal and prostate cancer patients. *PloS one* 2016: 11(1): e0146264.

158.   Ambrocio-Ortiz E, Pérez-Rubio G, Ramírez-Venegas A, Hernández-Zenteno RdJ, Paredes-López A, Sansores RH, Ramírez-Díaz ME, Cruz-Vicente F, Martínez-Gómez MdL, Falfán-Valencia R. Protective Role of Genetic Variants in HSP90 Genes-Complex in COPD Secondary to Biomass-Burning Smoke Exposure and Non-Severe COPD Forms in Tobacco Smoking Subjects. *Current Issues in Molecular Biology* 2021: 43(2): 887-899.

159.   Fritsch J, Fickers R, Klawitter J, Särchen V, Zingler P, Adam D, Janssen O, Krause E, Schütze S. TNF induced cleavage of HSP90 by cathepsin D potentiates apoptotic cell death. *Oncotarget* 2016: 7(46): 75774.

160.   Floen MJ, Forred BJ, Bloom EJ, Vitiello PF. Thioredoxin-1 redox signaling regulates cell survival in response to hyperoxia. *Free Radical Biology and Medicine* 2014: 75: 167-177.

161.   Márquez-Jurado S, Díaz-Colunga J, das Neves RP, Martinez-Lorente A, Almazán F, Guantes R, Iborra FJ. Mitochondrial levels determine variability in cell death by modulating apoptotic gene expression. *Nature communications* 2018: 9(1): 1-11.

162.   Dashdorj A, KR J, Lim S, Jo A, Nguyen MN, Ha J, Yoon K-S, Kim HJ, Park J-H, Murphy MP. Mitochondria-targeted antioxidant MitoQ ameliorates experimental mouse colitis by suppressing NLRP3 inflammasome-mediated inflammatory cytokines. *BMC medicine* 2013: 11(1): 1-13.

163.   Chae YC, Angelin A, Lisanti S, Kossenkov AV, Speicher KD, Wang H, Powers JF, Tischler AS, Pacak K, Fliedner S. Landscape of the mitochondrial Hsp90 metabolome in tumours. *Nature communications* 2013: 4(1): 1-10.

164.   Lassek WD, Gaulin SJ. Substantial but misunderstood human sexual dimorphism results mainly from sexual selection on males and natural selection on females. *Frontiers in Psychology* 2022: 13.

165.   Health NIo. NOT-OD-15-102: Consideration of Sex as a Biological Variable in NIH-funded Research. 2015. 2020.

166.   Caliman E, Petrella MC, Rossi V, Mazzoni F, Grosso AM, Fancelli S, Paglialunga L, Comin C, Roviello G, Pillozzi S. Gender matters. Sex-related differences in immunotherapy outcome in patients with non-small cell lung cancer. *Current Cancer Drug Targets* 2022.

167.   Tricarico R, Nicolas E, Hall MJ, Golemis EA. X-and Y-linked chromatin-modifying genes as regulators of sex-specific cancer incidence and prognosis. *Clinical Cancer Research* 2020.

168.   Han M, Murray S, Fell CD, Flaherty KR, Toews GB, Myers J, Colby TV, Travis WD, Kazerooni EA, Gross BH. Sex differences in physiological progression of idiopathic pulmonary fibrosis. *European Respiratory Journal* 2008: 31(6): 1183-1188.

169.   AIHW. Asthma Snapshot. 2020 [cited 22 November 2022]; Available from: https://www.aihw.gov.au/reports/chronic-respiratory-conditions/asthma/data

170.   AIHW. Chronic Obstructive Pulmonary Disease (COPD). 2020 [cited 24 November 2022]; Available from: https://www.aihw.gov.au/reports/chronic-respiratory-conditions/copd/contents/copd

171.   Tam A, Morrish D, Wadsworth S, Dorscheid D, Man SP, Sin DD. The role of female hormones on lung function in chronic lung diseases. *BMC women's health* 2011: 11(1): 24.

172.   Naeem A, Silveyra P. Sex Differences in Paediatric and Adult Asthma. *European Medical Journal (Chelmsford, England)* 2019: 4(2): 27.

173.   Reddy KD, Rutting S, Tonga K, Xenaki D, Simpson JL, McDonald VM, Plit M, Malouf M, Zakarya R, Oliver BG. Sexually dimorphic production of interleukin-6 in respiratory disease. *Physiological reports* 2020: 8(11): e14459.

174.   Balaton BP, Cotton AM, Brown CJ. Derivation of consensus inactivation status for X-linked genes from genome-wide studies. *Biology of sex differences* 2015: 6(1): 1-11.

175.   Wainer Katsir K, Linial M. Human genes escaping X-inactivation revealed by single cell expression data. *BMC genomics* 2019: 20(1): 1-17.

176.   Poleri C. Sex-Based Differences in Lung Cancer: Does It Matter? *Journal of Thoracic Oncology* 2022: 17(5): 599-601.

177.   Batdorj E, AlOgayil N, Zhuang QK, Galvez JH, Bauermeister K, Nagata K, Kimura T, Ward MA, Taketo T, Bourque G, Naumova AK. Genetic variation in the Y chromosome and sex-biased DNA methylation in somatic cells in the mouse. *Mammalian genome : official journal of the International Mammalian Genome Society* 2022.

178.   Godfrey AK, Naqvi S, Chmátal L, Chick JM, Mitchell RN, Gygi SP, Skaletsky H, Page DC. Quantitative analysis of Y-Chromosome gene expression across 36 human tissues. *Genome research* 2020: 30(6): 860-873.

179.   Andricovich J, Perkail S, Kai Y, Casasanta N, Peng W, Tzatsos A. Loss of KDM6A activates super-enhancers to induce gender-specific squamous-like pancreatic cancer and confers sensitivity to BET inhibitors. *Cancer cell* 2018: 33(3): 512-526. e518.

180.   Burgess JK, Boustany S, Moir LM, Weckmann M, Lau JY, Grafton K, Baraket M, Hansbro PM, Hansbro NG, Foster PS. Reduction of tumstatin in asthmatic airways contributes to angiogenesis, inflammation, and hyperresponsiveness. *American journal of respiratory and critical care medicine* 2010: 181(2): 106-115.

181.   Keatings VM, Collins PD, Scott DM, Barnes PJ. Differences in interleukin-8 and tumor necrosis factor-alpha in induced sputum from patients with chronic obstructive pulmonary disease or asthma. *American journal of respiratory and critical care medicine* 1996: 153(2): 530-534.

182.   Barnes PJ. The cytokine network in asthma and chronic obstructive pulmonary disease. *The Journal of clinical investigation* 2008: 118(11): 3546-3556.

183.   Rincon M, Irvin CG. Role of IL-6 in asthma and other inflammatory pulmonary diseases. *International journal of biological sciences* 2012: 8(9): 1281-1290.

184.   Liu T, Zhang L, Joo D, Sun S-C. NF-κB signaling in inflammation. *Signal transduction and targeted therapy* 2017: 2(1): 1-9.

185.   Kurai D, Saraya T, Ishii H, Takizawa H. Virus-induced exacerbations in asthma and COPD. *Frontiers in microbiology* 2013: 4: 293.

186.   Somayaji R, Chalmers JD. Just breathe: a review of sex and gender in chronic lung disease. *European Respiratory Review* 2022: 31(163).

187.   Stolz D, Kostikas K, Loefroth E, Fogel R, Gutzwiller FS, Conti V, Cao H, Clemens A. Differences in COPD exacerbation risk between women and men: analysis from the UK clinical practice research datalink data. *Chest* 2019: 156(4): 674-684.

188.   Jackson DJ, Johnston SL. The role of viruses in acute exacerbations of asthma. *Journal of Allergy and Clinical Immunology* 2010: 125(6): 1178-1187.

189.   Yuan X-h, Pang L-l, Yang J, Jin Y. Comparison of immune response to human rhinovirus C and respiratory syncytial virus in highly differentiated human airway epithelial cells. *Virology Journal* 2022: 19(1): 1-11.

190.   Huang S, He Q, Zhou L. T cell responses in respiratory viral infections and chronic obstructive pulmonary disease. *Chinese Medical Journal* 2021: 134(13): 1522-1534.

191.   Comer DM, Elborn JS, Ennis M. Comparison of nasal and bronchial epithelial cells obtained from patients with COPD. *PloS one* 2012: 7(3): e32924.

192.   Mihaylova VT, Kong Y, Fedorova O, Sharma L, Cruz CSD, Pyle AM, Iwasaki A, Foxman EF. Regional differences in airway epithelial cells reveal tradeoff between defense against oxidative stress and defense against rhinovirus. *Cell reports* 2018: 24(11): 3000-3007. e3003.

193.   Mitchell JE, Lund MM, Starmer J, Ge K, Magnuson T, Shpargel KB, Whitmire JK. UTX promotes CD8+ T cell-mediated antiviral defenses but reduces T cell durability. *Cell reports* 2021: 35(2).

194.   Cook KD, Shpargel KB, Starmer J, Whitfield-Larry F, Conley B, Allard DE, Rager JE, Fry RC, Davenport ML, Magnuson T. T follicular helper cell-dependent clearance of a persistent virus infection requires T cell expression of the histone demethylase UTX. *Immunity* 2015: 43(4): 703-714.

195.   Galaras A, Verykokakis M. UT (se) X differences during immune responses. *Nature Immunology* 2023: 24(5): 738-740.

196.   Busch SM, Lorenzana Z, Ryan AL. Implications for extracellular matrix interactions with human lung basal stem cells in lung development, disease, and airway modeling. *Frontiers in Pharmacology* 2021: 12: 645858.

197.   Cohen L, E X, Tarsi J, Ramkumar T, Horiuchi TK, Cochran R, DeMartino S, Schechtman KB, Hussain I, Holtzman MJ. Epithelial cell proliferation contributes to airway remodeling in severe asthma. *American journal of respiratory and critical care medicine* 2007: 176(2): 138-145.

198.   Barnes PJ. Small airway fibrosis in COPD. *The International Journal of Biochemistry & Cell Biology* 2019: 116: 105598.

199.   Jonkman JE, Cathcart JA, Xu F, Bartolini ME, Amon JE, Stevens KM, Colarusso P. An introduction to the wound healing assay using live-cell microscopy. *Cell adhesion & migration* 2014: 8(5): 440-451.

200.   Yasuda M, Harada N, Harada S, Ishimori A, Katsura Y, Itoigawa Y, Matsuno K, Makino F, Ito J, Ono J. Characterization of tenascin-C as a novel biomarker for asthma: utility of tenascin-C in combination with periostin or immunoglobulin E. *Allergy, Asthma & Clinical Immunology* 2018: 14: 1-12.

201.   Wu A, Zhang S, Liu J, Huang Y, Deng W, Shu G, Yin G. Integrated analysis of prognostic and immune associated integrin family in ovarian cancer. *Frontiers in genetics* 2020: 11: 705.

202.   Polosa R, Thomson NC. Smoking and asthma: dangerous liaisons. *European respiratory journal* 2013: 41(3): 716-726.

203.   Henson PM, Vandivier RW, Douglas IS. Cell death, remodeling, and repair in chronic obstructive pulmonary disease? *Proceedings of the American Thoracic Society* 2006: 3(8): 713-717.

204.   Bucchieri F, Marino Gammazza A, Pitruzzella A, Fucarino A, Farina F, Howarth P, Holgate ST, Zummo G, Davies DE. Cigarette smoke causes caspase-independent apoptosis of bronchial epithelial cells from asthmatic donors. *PloS one* 2015: 10(3): e0120510.

205.   Hardin M, Foreman M, Dransfield MT, Hansel N, Han MK, Cho MH, Bhatt SP, Ramsdell J, Lynch D, Curtis JL. Sex-specific features of emphysema among current and former smokers with COPD. *European Respiratory Journal* 2016: 47(1): 104-112.

206.   Yang CX, Shi H, Ding I, Milne S, Cordero AIH, Yang CWT, Kim EK-H, Hackett T-L, Leung J, Sin DD. Widespread sexual dimorphism in the transcriptome of human airway epithelium in response to smoking. *Scientific reports* 2019: 9(1): 1-14.

207.   Forsberg LA. Loss of chromosome Y (LOY) in blood cells is associated with increased risk for disease and mortality in aging men. *Human genetics* 2017: 136(5): 657-663.

208.   Gano JJ, Simon JA. A proteomic investigation of ligand-dependent HSP90 complexes reveals CHORDC1 as a novel ADP-dependent HSP90-interacting protein. *Molecular & Cellular Proteomics* 2010: 9(2): 255-270.

209.   Proud D, Leigh R. Epithelial cells and airway diseases. *Immunological reviews* 2011: 242(1): 186-204.

210.   Rinn JL, Snyder M. Sexual dimorphism in mammalian gene expression. *Trends in Genetics* 2005: 21(5): 298-305.

211.   Rao R, Chan SY. A dUTY to Protect: Addressing "Y" We See Sex Differences in Pulmonary Hypertension. *American Journal of Respiratory and Critical Care Medicine* 2022(ja).

212.   Barnes P. Similarities and differences in inflammatory mechanisms of asthma and COPD. *Breathe* 2011: 7(3): 229-238.

213.   Dyer DP. Understanding the mechanisms that facilitate specificity, not redundancy, of chemokine-mediated leukocyte recruitment. *Immunology* 2020: 160(4): 336-344.

214.   Radzikowska U, Golebski K. Sex hormones and asthma: the role of estrogen in asthma development and severity. Wiley Online Library, 2022.

215.   Keselman A, Heller N. Estrogen signaling modulates allergic inflammation and contributes to sex differences in asthma. *Frontiers in immunology* 2015: 6: 568.

216.   Bonds RS, Midoro-Horiuti T. Estrogen effects in allergy and asthma. *Current opinion in allergy and clinical immunology* 2013: 13(1): 92.

217.   Reyes-García J, Montaño LM, Carbajal-García A, Wang Y-X. Sex Hormones and Lung Inflammation. Lung Inflammation in Health and Disease, Volume II. Springer, 2021; pp. 259-321.

218.   Chiarella SE, Cardet JC, Prakash Y. Sex, Cells, and Asthma. In: Mayo Clinic Proceedings; 2021: Elsevier; 2021. p. 1955-1969.

219.   Fuentes N, Silveyra P. Endocrine regulation of lung disease and inflammation. *Experimental Biology and Medicine* 2018: 243(17-18): 1313-1322.

220.   Karadag F, Ozcan H, Karul A, Yilmaz M, Cildag O. Sex hormone alterations and systemic inflammation in chronic obstructive pulmonary disease. *International journal of clinical practice* 2009: 63(2): 275-281.

221.   Xie G, Liu X, Zhang Y, Li W, Liu S, Chen Z, Xu B, Yang J, He L, Zhang Z. UTX promotes hormonally responsive breast carcinogenesis through feed-forward transcription regulation with estrogen receptor. *Oncogene* 2017: 36(39): 5497-5511.

222.   Borkar NA, Combs CK, Sathish V. Sex Steroids Effects on Asthma: A Network Perspective of Immune and Airway Cells. *Cells* 2022: 11(14): 2238.

223.   Kalra N, Ishmael FT. Cross-talk between vitamin D, estrogen and corticosteroids in glucocorticoid resistant asthma. *OA inflammation* 2014: 2(1): 2.

224.   Galan-Caridad JM, Harel S, Arenzana TL, Hou ZE, Doetsch FK, Mirny LA, Reizis B. Zfx controls the self-renewal of embryonic and hematopoietic stem cells. *Cell* 2007: 129(2): 345-357.

225.   Arnold AP, Chen X. What does the "four core genotypes" mouse model tell us about sex differences in the brain and other tissues? *Frontiers in neuroendocrinology* 2009: 30(1): 1-9.