Role of epigenetic mechanisms in the pathogenesis of chronic respiratory diseases and response to inhaled exposures: from basic concepts to clinical applications.

Renata Z. Jurkowska*

Division of Biomedicine, School of Biosciences, Cardiff University, Cardiff, UK

*Correspondence should be addressed to jurkowskar@cardiff.ac.uk, Telephone +44 29 2087 9067

Abstract

Epigenetic modifications are chemical groups in our DNA (and chromatin) that determine which genes are active and which are shut off. Importantly, they integrate environmental signals to direct cellular function. Upon chronic environmental exposures, the epigenetic signature of lung cells gets altered, triggering aberrant gene expression programs that can lead to the development of chronic lung diseases. In addition to driving disease, epigenetic marks can serve as attractive lung disease biomarkers, due to early onset, disease specificity, and stability, warranting the need for more epigenetic research in the lung field.

Despite substantial progress in mapping epigenetic alterations (mostly DNA methylation) in chronic lung diseases, the molecular mechanisms leading to their establishment are largely unknown. This review is meant as a guide for clinicians and lung researchers interested in epigenetic regulation with a focus on DNA methylation. It provides a short introduction to the main epigenetic mechanisms (DNA methylation, histone modifications and non-coding RNA) and the machinery responsible for their establishment and removal. It presents examples of epigenetic dysregulation across a spectrum of chronic lung diseases and discusses the current state of epigenetic therapies. Finally, it introduces the concept of epigenetic editing, an exciting novel approach to dissecting the functional role of epigenetic modifications. The promise of this emerging technology for the functional study of epigenetic mechanisms in cells and its potential future use in the clinic is further discussed.

Keywords: epigenetic modifications, chronic lung diseases, DNA methylation, epigenetic biomarkers, epigenetic therapies, epigenetic editing

List of abbreviations:

AHRR - aryl hydrocarbon receptor repressor

AM – alveolar macrophages

ATAC-seq – assay for transposase-accessible chromatin with sequencing

BAL – bronchoalveolar lavage

dCas9 – catalytically inactive / dead Cas 9 nuclease (from the CRISPR system)

ChIP-seq – chromatin immunoprecipitation with sequencing

COPD - chronic obstructive pulmonary disease

CRISPR - Clustered Regularly Interspaced Short Palindromic Repeats

CpG site – cytosine – guanine dinucleotide separated by a phosphate

DNMT - DNA methyltransferase

DNMTi – DNA methyltransferase inhibitor

EWAS - epigenome-wide association studies

FFPE - formalin-fixed paraffin-embedded

HDAC – histone deacetylase

IPF – idiopathic pulmonary disease

lncRNAs - long non-coding RNAs

miRNA – micro-RNA

ncRNAs – non-coding RNAs

MBD – methyl binding domain

MTase – methyltransferase

NGS – next-generation sequencing

PTM – posttranslational modification

RRBS - reduced-representation bisulfite sequencing

SAM - S-adenosyl-L-methionine

scRNA-seq- single-cell RNA sequencing

TET proteins - Ten-Eleven-Translocation proteins

TDG - Thymine DNA glycosylase

TF – transcription factor

UHRF1 - ubiquitin-like, containing PHD and RING finger domains 1

WGBS – whole genome bisulfite sequencing

5mC – 5-methylcytosine

5hmC – 5-hydroxymethylcytosine

5fC – 5-formylcytosine

5caC – 5-carboxylcytosine

Table of Contents

1. Epigenetics – an ingenious mechanism for regulating genetic information.

The advent of the next-generation sequencing (NGS) technologies revolutionised biomedical research. We deciphered the sequence of the human genome (Lander, et al., 2001; Nurk, et al., 2022; Venter, et al., 2001) and identified genetic drivers of numerous human diseases (including lung diseases), bringing biomedicine into the genomic era (Ashley, 2016). The development of single-cell transcriptomic technologies offered further unprecedented insights into the complexity and cellular identity of the human and mouse lungs (Negretti, et al., 2021; Travaglini, et al., 2020). It led to the identification of novel cell types (Montoro, et al., 2018; Plasschaert, et al., 2018) and differentiation states (Ruiz Garcia, et al., 2019). It revealed unique aspects of disease processes in the human lung, opening doors to the functional dissection of cellular pathology, and moving the research into the single-cell transcriptomic era (Adams, et al., 2023; Van de Sande, et al., 2023). However, our understanding of the molecular processes orchestrating gene expression and cellular differentiation programs in the human lung is still very limited. Similarly, the regulatory mechanisms driving lung cell dysfunction in response to environmental stimuli are not well understood. To provide the missing link between the environment, our genes, and their regulation in disease we need to move lung research into the epigenetic era. For further progress and translation of epigenetic research to the clinics, there is an urgent need for functional studies on how epigenetic mechanisms in lung cells are impacted by environmental insults and how their dysregulation drives the development of chronic lung disease.

We begin as a single cell, the zygote, and yet there are more than 400 cell types in the adult human body (Hatton, et al., 2023). How can a single genetic sequence give rise to so many different cells, each with unique morphology, gene expression programs and function? We know now that lineage-specific gene expression programmes are initiated by transcription

factors and modulated by epigenetic mechanisms that regulate how the genetic information is read and interpreted by the cellular machinery. The development of epigenetic systems was proposed as a fundamental step in the evolution of multicellular organisms because of their need to maintain stable cellular differentiation (Jeltsch, 2013). Epigenetic mechanisms are defined as “heritable changes in gene function that cannot be explained by changes in the DNA sequence” (Riggs, 1996).

To fit in the cell nucleus, our DNA is tightly packaged into chromatin. The basic building unit of the chromatin, the nucleosome, consists of DNA wrapped around an octamer of histone proteins (two of each of the four core histones H2A, H2B, H3 and H4). Both the DNA and the histone proteins (especially their tails that protrude from the nucleosome surface) are decorated with a variety of chemical groups, called epigenetic modifications. This epigenetic code determines which genes are active and which are silenced. The regulatory function of epigenetic modifications is mediated by controlling the accessibility of the DNA to the cellular machinery and regulating the binding of transcriptional activators and repressors (Allis & Jenuwein, 2016). The main epigenetic mechanisms include DNA methylation, posttranslational modification (PTM) of histones, chromatin remodelling as well as non-coding RNA (ncRNA) (**Figure 1**).

In this review, I will briefly introduce the main epigenetic mechanisms and examine their dysregulation in a variety of chronic lung diseases, with a focus on DNA methylation. I will present the main technologies used for mapping and functional validation of DNA methylation alterations and discuss the promise of epigenetic editing and epigenetic therapies.

2. DNA methylation – a small group with a big biological impact.

DNA methylation is the best-described example of an epigenetic mechanism and a key epigenetic signal used for the direct control of gene expression [for general reviews on DNA methylation, please refer to (Greenberg & Bourc'his, 2019; Jurkowska, Jurkowski, et al., 2011; Schubeler, 2015; Smith, et al., 2024; Smith & Meissner, 2013)]. Owing to this vital regulatory function, DNA methylation plays important biological roles in regulating genome stability (through silencing repetitive elements), genomic imprinting, X-chromosome inactivation and regulation of gene expression in development and response to environmental and occupational cues. The importance of DNA methylation is demonstrated by the lethal phenotypes of the genetic knockouts of any of the active DNA methyltransferase enzymes in mice (Li, et al., 1992; Okano, et al., 1999) and by the ever-growing number of diseases associated with altered DNA methylation signatures, including pulmonary and non-pulmonary disease conditions (Bergman & Cedar, 2013; R. Liu, et al., 2023; Michalak, et al., 2019; Zhao, et al., 2021).

The methylation of human and mammalian DNA was identified in the early 80s [reviewed in (Jurkowska & Jeltsch, 2022b)]. It arises through the covalent addition of the methyl group (CH_3) to the cytosine residues in the DNA, generating 5-methylcytosines (5mC) (**Figure 2**). In humans, it occurs predominantly in the context of CpG sites, where a cytosine is followed by a guanine residue. Non-CpG DNA methylation (arising on CA, CC or CT sites) occurs at much lower levels in differentiated tissues and is mostly present in embryonic stem cells and the brain (Schultz, et al., 2015).

Depending on the cell type, roughly 60-80% of all CpG sites in the DNA are methylated, corresponding to 3-8% of all cytosines in the human genome (Greenberg & Bourc'his, 2019).

Importantly, only selected CpG sites are methylated, resulting in the generation of a tissue and cell-type-specific pattern consisting of methylated and non-methylated sites. Hence, the DNA methylation pattern confers the epigenetic cell identity (Loyfer, et al., 2023). This has important consequences for studying epigenetic modifications in a complex organ or tissue (like for example lung tissue, epithelium or blood), as different cell types will carry unique patterns of DNA methylation (and other epigenetic modifications). Consequently, the DNA methylation profile (the methylome) from bulk tissues represents an average of different cell types, limiting our understanding of cell-type-specific contributions to disease development. Thus, studies of purified lung cell populations or single-cell approaches are urgently needed to investigate the role of DNA methylation (and other epigenetic modifications) in driving cellular function in healthy and diseased lungs.

Regions of the genome characterised by a high density of CpG sites are called CpG islands. They are found in the promoters (or the first exons) of around two-thirds of human genes, and less commonly in gene bodies. They constitute regulatory units for DNA methylation, as their methylation correlates with the transcriptional activity of their neighbouring genes (Jones, 2012). In healthy somatic tissues, CpG islands in gene promotors are usually unmethylated, whereas CpG islands located in gene bodies tend to become methylated during development depending on the transcriptional activity of their genes (Jeziorska, et al., 2017). The aberrant DNA methylation of CpG islands in the promoters of tumour suppressor genes, leading to their inactivation, is a key hallmark of epigenetic dysregulation in cancer cells (Bergman & Cedar, 2013).

How can a small methyl group exert a biological effect? The methyl group of the methylated cytosine is positioned in the major groove of the DNA, where it can be recognised by proteins

interacting with DNA, for example, transcription factors. DNA methylation can regulate their binding to the DNA and thereby modulate gene expression (Yin, et al., 2017). In addition, DNA methylation can recruit specific methyl reader proteins (**Figure 1**) (e.g., so-called methyl-binding domain (MBD) proteins) or additional epigenetic modifiers (e.g. histone-modifying enzymes or chromatin remodelling enzymes), leading to changes in the accessibility of chromatin to the transcriptional machinery. The exact mechanism by which DNA methylation affects chromatin structure is not fully understood, but it is known that methylated DNA is closely associated with a closed, relatively inactive chromatin structure (Keshet, et al., 1986). As DNA methylation adds additional information to the genetic code that is not encoded in the DNA sequence itself and is heritable through cell divisions, the 5mC is sometimes referred to as the "5th letter of the genetic alphabet". In addition to the 5-methylcytosine, 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (**Figure 2**), which arise in the process of active DNA demethylation, are present in the human DNA, further extending the (epi)genetic alphabet.

In the early 1980s, global loss of DNA methylation (DNA hypomethylation) was reported in cancer cells (Feinberg & Vogelstein, 1983; Gama-Sosa, et al., 1983). This finding was followed by the identification of the aberrant gain of methylation (DNA hypermethylation) in the promoters of tumour suppressor genes, leading to their inactivation in cancer (Baylin, et al., 1986; Greger, et al., 1989). These seminal discoveries exemplified DNA methylation changes characteristic of most cancer cells and demonstrated the fundamental role of DNA methylation in disease development and progression (Bergman & Cedar, 2013). Since these initial discoveries, the number of diseases associated with altered DNA methylation signatures in different organs has been constantly growing (Bergman & Cedar, 2013; R. Liu, et al., 2023; Michalak, et al., 2019; Zhao, et al., 2021).

2.1. Molecular machinery setting and erasing DNA methylation.

The methyl groups are added to the DNA by specific enzymes called DNA methyltransferases (MTases or DNMTs) and can be removed by the Ten-Eleven-Translocation (TET) DNA demethylases (Jurkowska & Jeltsch, 2022a; Jurkowska & Jurkowski, 2019; Ravichandran, et al., 2018). Three active MTases (DNMT1, DNMT3A and DNMT3B) and three TET enzymes (TET1, TET2 and TET3) exist in humans (and other mammals). DNMT3A and DNMT3B proteins (with the help of a stimulatory factor DNMT3L) introduce DNA methylation patterns during early embryonic development and gametogenesis (**Figure 3**). They are highly expressed in undifferentiated cells and germ cells, and present at much lower levels in somatic differentiated cells. After establishment, patterns of DNA methylation are mostly preserved, with only small tissue-specific changes. However, they can get significantly altered in response to environmental exposures and in diseases (as discussed below).

DNA methylation information consisting of methylated and unmethylated sites is preserved over rounds of cell divisions by a well-designed maintenance mechanism, which enables the function of DNA methylation as a key epigenetic mechanism mediating long-term gene repression. As CpG sites are symmetric and usually methylated in both strands of DNA, the methylation signal is present in two DNA strands. The process of DNA replication generates a strand of unmethylated DNA, converting fully methylated CpGs into hemimethylated sites (with one strand unmethylated and one methylated) (**Figure 3**). These sites are then re-methylated by a maintenance methyltransferase enzyme (DNMT1), a molecular copy machine, which is highly expressed in proliferating cells, localizes to the replication forks and has a high preference towards hemimethylated DNA (Jeltsch & Jurkowska, 2014; Petryk, et al., 2021). The recruitment of DNMT1 to hemimethylated DNA during replication is facilitated by a chromatin factor, UHRF1 (ubiquitin-like, containing PHD and RING finger domains 1), a key

component of the DNA maintenance machinery and regulator of the DNMT1 enzyme (Bostick, et al., 2007).

DNA methylation was once believed to be an irreversible epigenetic signal. Despite the existence of a specific maintenance mechanism, DNA methylation can be lost by either a passive mechanism, when the maintenance MTase activity is absent over consecutive cycles of cell divisions or via an active DNA demethylation process (**Figure 3**). Active DNA demethylation involves stepwise oxidation of 5mC to 5hmC, 5fC and 5caC by TET enzymes (Ito, et al., 2011) (**Figure 2**). The oxidised bases are then removed by the Thymine DNA glycosylase (TDG) and the base excision repair (BER) mechanism, restoring DNA to the unmethylated state (Jurkowska & Jurkowski, 2019; Ravichandran, et al., 2018; X. Wu & Zhang, 2017). TET1 and TET2 are expressed in multiple tissues, including embryonic stem cells, but are generally downregulated during differentiation, whereas TET3 is mostly present in oocytes, the zygote, and neurons [reviewed in (Jurkowska & Jurkowski, 2019; X. Wu & Zhang, 2017)].

Hence, the patterns of DNA methylation in living cells are shaped by the combined action of DNA methyltransferases and TET demethylases and their targeting, localization and activity need to be precisely controlled.

2.2. Domain structure of human methyltransferase and demethylases

Human DNA methyltransferases and demethylases are all large, multi-domain proteins. They contain a large regulatory N-terminal part and a smaller C-terminal part involved in catalysis [reviewed in (Jurkowska & Jeltsch, 2022a)]. The N-terminal part contains several domains that mediate the localization of the enzymes to the cell nucleus and regulate their interaction with other proteins, chromatin and DNA. The C-terminal domain of DNMTs is required for the

binding of the methylation cofactor (S-adenosyl-L-methionine, SAM), recognition of the DNA and catalysis. The catalytic core of TET enzymes contains a conserved double-stranded β-helix (DSBH) domain, a cysteine-rich domain, and binding sites for the substrate DNA, as well as iron (Fe 2+) and 2-oxoglutarate (2-OG), two key cofactors required for catalysis (Ravichandran, et al., 2018).

Notably, the spatial arrangement of the various domains in DNMTs plays a crucial role in the regulation of their function, providing an elegant mechanism for direct control of the enzymes' activity and function in cells [reviewed in (Jeltsch & Jurkowska, 2016)]. Interaction with protein partners (e.g. stimulatory factor DNMT3L), other chromatin modifications (e.g. specific PTMs on histone tails) or modifications of the linkers between various domains can lead to rearrangements of the DNMTs domains leading to their allosteric activation or inhibition (Bashtrykov, et al., 2014; Rajavelu, et al., 2018).

Unlike most transcription factors, DNMTs and TET enzymes do not display strict sequence specificity beyond the preference for the CpG sites and can modify CpG sites in a variety of sequence contexts. However, they are sensitive to the sequences flanking their target sites (Adam, et al., 2022; Handa & Jeltsch, 2005; Jurkowska, Siddique, et al., 2011) and display up to 250-fold preference for the most favorable sites (Ravichandran, et al., 2018). In addition, DNMT3 MTases form heterotetrameric complexes with a stimulatory factor DNMT3L (Jia, et al., 2007). The arrangement in the complex positions the active sites at e specific distance, allowing simultaneous methylation of CpGs separated by 8–10 bps (Jurkowska, Rajavelu, et al., 2011). These intrinsic biochemical properties of DNMTs and TETs may contribute to DNA target selection and the establishment of DNA methylation patterns in development but also

in disease. Indeed, the imprints of these properties have been observed in genome-wide DNA methylation studies (Jeltsch & Jurkowska, 2013; Ravichandran, et al., 2022).

Despite significant progress in understanding the biochemical properties and the mechanism of the human DNA methyltransferases and demethylases, their genomic targeting as well as their regulation in cells is not well understood. Even less is known about the expression and regulation of DNMTs and TET enzymes in healthy and diseased lungs. This is a significant research gap that should be addressed, as the number lung diseases with altered methylation signatures is constantly growing and epigenetic enzymes and pathways constitute attractive druggable targets for clinical applications (as described below).

3. Histone posttranslational modifications – sophisticated code for gene regulation.

In addition to the DNA, all four core histone proteins in the nucleosomes (H2A, H2B, H3 and H4) and the linker histone H1 can be subject to posttranslational modifications (PTMs), including acetylation, phosphorylation, methylation, ubiquitination and sumoylation, among others. Less-studied examples of histone PTMs include ADP ribosylation, deimination, proline isomerization, crotonylation, propionylation, butyrylation, formylation, hydroxylation and O-GlcNAcylation [for a recent comprehensive review of different PTMs, please refer to (Millán-Zambrano, et al., 2022)]. These modifications occur on specific amino acid residues of the histones, mostly on their N-terminal flexible tails that protrude from the nucleosome core (**Figure 4**). Histone PTMs are introduced and removed by specific classes of enzymes (writers and erasers, respectively), which often reside in large protein complexes carrying multiple enzymatic activities. The pattern of histone PTMs constitutes a sophisticated molecular code that can be specifically recognised and interpreted by a variety of proteins (so-called epigenetic readers) that influence chromatin structure and thus gene expression (Turner,

1993). In addition to recruiting protein readers, PTMs can also directly modify the physical properties of the histone tails and thereby directly impact chromatin accessibility. Histone acetylation and phosphorylation are examples of PTMs that can directly influence nucleosome packaging through the modification of charges.

Specific patterns of histone modifications have been suggested to provide activating ("ON") or repressing ("OFF") signatures for gene expression. For example, histone acetylation often correlates with transcriptionally active, open chromatin, which is devoid of DNA methylation whereas deacetylated histones are found in repressed, condensed regions, which also contain DNA methylation (**Figure 5**)(Cosgrove, et al., 2004). However, the same type of modification (e.g. methylation of lysines) may have opposite effects on gene expression depending on the context, localisation and the exact residue that is modified. For example, trimethylation of lysine 9 of histone H3 (H3K9me3) is associated with repressed chromatin, trimethylation of lysine 4 on H3 (H3K4me3) marks active promoters, while monomethylation of lysine 4 on H3 (H3K4me1) is associated with enhancer elements. The complexity of the histone code is further increased by the crosstalk between different adjacent modifications or even between different histone tails (Millán-Zambrano, et al., 2022).

Thanks to the multi-institutional efforts of international epigenomic initiatives (like for example, NIH Roadmap Epigenomics (Bernstein, et al., 2010), Encyclopaedia of DNA Elements (ENCODE) project (Consortium, 2004), Blueprint project (Martens & Stunnenberg, 2013) or the International Human Epigenome Consortium (IHEC) (Stunnenberg, et al., 2016)), various histone modifications have been mapped at high resolution across numerous human tissues and cells, providing reference epigenomic maps. These efforts led to the identification of chromatin states characterised by specific types of histones PTMs and provided key evidence

of the association of chromatin modifications with the regulation of gene expression. They can be used to define different functional elements of the human genome (e.g. promoters, proximal and distal enhancers, repressed regions etc) and enable the integration of multiple omics data.

4. Non-coding RNA.

Even though a large fraction of our genetic information is transcribed into RNA, only a low percent of these transcripts encodes actual proteins. The majority of all RNAs are non-coding RNA (ncRNA) and their function remains an area of intense research. Non-coding RNAs contribute another key component to epigenetic regulation in living cells [for recent reviews, please refer to (Mattick, et al., 2023; Nemeth, et al., 2024)]. They are classified by size into long ncRNA (lncRNAs, >200 bps) and short ncRNAs (<200 bps), and play prominent roles in the regulation of transcription, silencing of transposons, and RNA modification, among others. RNA research is evolving very fast and new classes of ncRNAs, e.g. enhancer RNAs or circular RNAs have been recently identified (Uszczynska-Ratajczak, et al., 2018). Long non-coding RNAs play a prominent role in the recruitment of various chromatin factors to their target sites, leading to the establishment of silenced or active chromatin domains (**Figure 1**). The most remarkable example, where a single lncRNA can trigger epigenetic silencing of an entire chromosome is the X inactive-specific transcript (XIST) RNA, which is critical for the inactivation of X-chromosome and dosage compensation in females (Loda & Heard, 2019). The best-studied examples of small ncRNAs are micro RNAs (miRNAs), which are endogenous small (between 19 and 22 nucleotides in length) ncRNAs. They play important roles in the regulation of gene expression by controlling mRNA translation (Nemeth, et al., 2024).

5. Reader domains confer biological functions of epigenetic modifications.

One of the interesting questions that has been keeping the epigenetic field busy is how the message encoded in the pattern of modifications of DNA and histones is read by the cellular machinery to exert biological function. This interest led to the discovery of so-called epigenetic readers, protein domains that can specifically recognise and bind to defined epigenetic modifications and trigger downstream signalling (**Figure 1**). Bromodomains, which recognise acetylated lysines and chromodomains, which bind to methylated lysines are two best-known examples of such reading domains in humans (Franklin, et al., 2022). In addition, the methylated cytosines on DNA can also be recognised by methyl-binding domain (MBDs) proteins that transmit the downstream signalling generally associated with transcriptional repression (**Figure 1**). In addition to the epigenetic enzymes that set the epigenetic mark (writers) and remove it (erasers), epigenetic readers are also an attractive class of druggable targets for clinical applications (as discussed below).

6. Methods for genome-wide mapping of DNA methylation.

Because epigenetic mechanisms, like DNA methylation, provide an interface between environmental factors and the genome, the methylation pattern of a cell or tissue can also be considered a cellular memory of past exposures. This is particularly relevant to the human lung, which is constantly exposed to environmental insults. Hence, mapping DNA methylation (and other epigenetic modification) changes across the course of a disease may lead to the identification of molecular pathways driving its onset and progression. In addition to driving disease development, DNA methylation can serve as attractive disease biomarkers, due to their early onset, disease specificity, and stability (Costa-Pinheiro, et al., 2015) and can be used to develop DNA methylation-based predictors of health and disease (Yousefi, et al., 2022).

With the growing recognition of the importance of DNA methylation in regulating human health, arose a need for the development of robust methods for the detection of DNA methylation in a genome-wide manner.

The most popular use of DNA methylation in epidemiology is in epigenome-wide association studies (EWAS), which are population-based studies that analyse the association between DNA methylation levels (or any other epigenetic modification) and a specific exposure (for example, cigarette smoking or air pollution), disease or phenotype (Birney, et al., 2016).

Several methods can be employed to measure DNA methylation at thousands or millions of CpG sites simultaneously [for a general review on methods of DNA methylation, refer to (Laird, 2010)]. Genome-wide interrogation of DNA methylation at single CpG sites can be performed via array technologies or by harnessing the power of next-generation sequencing (NGS) (**Figure 6**). The gold standard in the field relies on the chemical treatment of the DNA with sodium bisulfite which enables the differentiation of unmethylated from methylated cytosines. This is possible because, upon bisulfite treatment, unmethylated cytosines are converted to uracil, whereas methylated cytosines are protected (**Figure 6**). The bisulfite-based methods offer single-nucleotide resolution and require low input of DNA, making them applicable to small amounts of human material. Of note, classical bisulfite sequencing cannot distinguish between 5mC and 5hmC, therefore, data reported as DNA methylation will likely be a mix of DNA methylation and hydroxymethylation. This is important, as emerging research indicates that 5hmC is not just intermediate in DNA demethylation, but a novel epigenetic modification with regulatory functions on its own (Bachman, et al., 2014). Several new techniques have now been developed to differentiate between the two epigenetic states and allow reliable mapping of all four modified cytosine bases (5mC, 5hmC, 5fC and 5caC) in the

human DNA (Booth, et al., 2023; Y. Liu, et al., 2021; Y. Liu, et al., 2019; Schutsky, et al., 2018; Vaisvila, et al., 2021; M. Yu, et al., 2012).

6.1. DNA Methylation arrays.

Due to their robust performance, cost-effectiveness, scalability and user-friendly data analysis, Illumina DNA methylation arrays have been widely used to investigate DNA methylation patterns in large cohorts in disease research and are a method of choice for EWAS in epidemiology. Different arrays have been developed in the past 15 years, the most well-known are the human BeadChip arrays which are based on Illumina's Infinium technology. They enable the hybridization of DNA to specific probes immobilised on beads (hence the name BeadChip) that distinguish methylated and unmethylated sequenced based on their differential sequence after bisulfite treatment. They have evolved over several generations, starting from 27K (Bibikova, et al., 2009), 450K (Bibikova, et al., 2011), EPIC (Pidsley, et al., 2016), to the most recent EPIC version 2 (EPICv2) (D. Kaur, et al., 2023). Each generation extended the array's coverage of the human genome and showed improved probe design.

The first one, the 27K BeadChip array (Bibikova, et al., 2009) included around 27,000 probes covering CpG sites located within promoter regions, focusing on genes implicated in cancer. Infinium HumanMethylation450 Beadchip (also known as the 450K array) which measures over 450,000 CpG sites, included representation of gene bodies (Bibikova, et al., 2011), offering greatly improved genomic coverage. The EPIC array (also known as the 850K or EPICv1), released in 2015, expanded the coverage of cis-regulatory elements identified by the ENCODE (Consortium, 2012) and FANTOM5 (Lizio, et al., 2015) projects, enabling the analysis of numerous enhancers (Pidsley, et al., 2016). The latest addition, the Infinium BeadChip

EPICv2, with further expanded genomic coverage (>935,000 CpG sites), was recently launched by Illumina.

The versality of the DNA methylation arrays is demonstrated by their compatibility with formalin-fixed paraffin-embedded (FFPE) samples (Moran, et al., 2014), enabling analysis of DNA methylation patterns from archived collections. It is further enhanced by the ability to examine other cytosine modifications, like the recently identified 5-hydroxymethylcytosine mark (Stewart, et al., 2015). However, despite increased overall coverage, due to the cell type specificity of the distal regulatory elements (like enhancers), their coverage for any cell type on the arrays remains limited. Recently, some concerns have been raised regarding low concordance across 450K and EPIC platforms in different tissues (Olstad, et al., 2022; Sugden, et al., 2020; Zhuang, et al., 2024), making comparing data and replication challenging and illustrating a need for better cross-platform translatability. Measuring other epigenetic marks, such as histone modifications in hundreds of clinical samples, with a high degree of accuracy is not yet possible and necessitates further technological developments.

6.2. Genome-wide approaches.

While cost-effective and thus applicable to larger cohorts, DNA methylation arrays only allow interrogation of the selected parts of the genome, covering less than <1 million out of the 28 million of the CpG sites present in the human genome. Whole genome bisulfite sequencing (WGBS) offers the most comprehensive coverage. It harnesses the power of next-generation sequencing, leveraging single-nucleotide resolution with genome-wide coverage (Singer, 2019) (**Figure 6**). Its higher sequencing cost makes it more suitable for smaller-scale discovery research, complementing array-based larger-cohort initiatives. Alternatively, reduced-representation bisulfite sequencing (RRBS), which includes enrichment of regulatory regions,

combines very good coverage (around 2 million CpG sites) with lower sequencing costs (Meissner, et al., 2005). It uses the digestion of genomic DNA with restriction enzymes followed by size selection of the DNA to focus the analysis on the genomic regions containing CpG dinucleotides, which are the main targets of DNA methylation.

Sequencing-based approaches have an additional advantage over arrays, as they provide a binary read-out of the methylation state of individual CpG sites (which can be either methylated or unmethylated) that is easier to interpret than probe signal intensities of the DNA methylation arrays. With the lowering cost of sequencing, they are predicted to replace array technologies in the future (Yousefi, et al., 2022). The development of reliable methods for quantification of DNA methylation on a genome-wide scale has enabled fine mapping of DNA methylation dysregulation across various lung diseases, empowering investigation of epigenetic mechanisms and future biomarker development (**Figure 6**).

7. Epigenetic dysregulation in chronic lung diseases.

As chronic lung conditions are largely caused by environmental exposures (in genetically susceptible individuals), which are known to remodel the epigenetic landscape of cells; it is not surprising that genome-wide aberrant DNA methylation changes have been identified in response to smoking, air pollution and across a spectrum of chronic lung diseases.

7.1. Cigarette smoking profoundly remodels the epigenome of cells.

Cigarette smoking is the best-described example of an environmental exposure with a well-established massive impact on the epigenetic landscape in humans. Selected examples are discussed below, but the reader is encouraged to refer to (Gao, et al., 2015; G. Kaur, et al., 2019; Silva & Kamens, 2021; Zong, et al., 2019) for recent systematic reviews on the association of smoking exposure with epigenetic changes.

Smoking is a major cause of premature death worldwide (Ezzati & Lopez, 2003) and a risk factor for the development of several human diseases, including chronic obstructive pulmonary disease (COPD), Idiopathic pulmonary disease (IPF), several cancers (including lung cancer), cardiovascular disease, osteoporosis, and others (Prevention, 2014; Wiklund, et al., 2019). Notably, even after smoking cessation, the risk for some diseases persists, indicating the existence of a cellular memory of smoking exposure. Due to its biochemical stability and heritability through cellular divisions, DNA methylation emerged as one of the mechanisms that may confer this long-term risk. Correspondingly, several studies indicated that DNA methylation changes from prenatal exposure to maternal smoking may persist into adulthood (Hoang, et al., 2024; Richmond, et al., 2015) and confer future risk of disease (Wiklund, et al., 2019).

Multiple studies consistently demonstrated genome-wide alterations in DNA methylation in smokers (**Table 1**), indicating a broad remodelling of the epigenome in response to cigarette smoke exposure in the blood (Chen, et al., 2013; Shenker, et al., 2013b; Zeilinger, et al., 2013), lung tissue (Bosse, et al., 2012; Seiler, et al., 2020), sputum (Belinsky, et al., 2002), buccal mucosa (Wan, et al., 2015) and airway epithelium (Buro-Auriemma, et al., 2013; F. Liu, et al., 2010). In addition, statistically robust CpG sites associated with various smoking-related parameters, comprising smoking status (current, never, or former) (Joehanes, et al., 2016), time since smoking cessation (Zeilinger, et al., 2013), as well as cumulative exposures (Guida, et al., 2015; McCartney, Stevenson, et al., 2018) were identified. These studies led to the development of DNA methylation-based predictors for smoking behaviours (Bollepalli, et al., 2019; Maas, et al., 2019; McCartney, Hillary, et al., 2018; McCartney, Stevenson, et al., 2018), which may prove an attractive alternative to self-reported smoking or current smoking

biomarkers (e.g., serum cotinine measurements) in clinical practice (Bojesen, et al., 2017). These studies exemplify the usefulness of DNA methylation as a biomarker of exposure.

One of the largest EWAS in adults on cigarette smoking to date included 15 907 participants from 16 different cohorts. Using Illumina 450K BeadChip array, the authors identified differential DNA methylation between current and non-smokers at more than 2500 CpG sites mapped to 1405 genes (Joehanes, et al., 2016), highlighting the profound impact of smoking on the epigenetic landscape of blood cells. Notably, genes with smoke-induced DNA methylation changes were also enriched for genes associated with smoking-related diseases, like COPD or cardiovascular disease in genome-wide association studies (GWAS). Enrichment of sites with altered DNA methylation in gene regulatory regions, like island shores, gene bodies, and enhancers suggested that smoking-associated changes may impact gene expression. However, this association was not functionally investigated (Joehanes, et al., 2016).

A very recent large (5 cohorts, 15,014 participants) meta-analysis of smoking using a more comprehensive Illumina EPIC array replicated many of the previously reported associations and identified an additional 1405 genes with methylation changes in CpG sites not covered by the 450K array (Hoang, et al., 2024).

7.1.1. The curious case of the AHRR gene.

The most prominent and best-characterised example of smoking-associated DNA methylation alterations occurs in the human AHRR (aryl hydrocarbon receptor (AHR) repressor) gene. AHRR is a transcription factor repressing the aryl hydrocarbon receptor-dependent gene expression. The AHR signalling cascade, which mediates dioxin toxicity, is involved in the detoxification of compounds from tobacco smoke (like polyaromatic hydrocarbons) and is

involved in the regulation of cell growth, apoptosis and differentiation and the modulation of the immune system (Vogel & Haarmann-Stemmann, 2017).

In the AHRR gene, more than 100 differentially methylated sites associated with various smoking-related parameters were identified (Silva & Kamens, 2021; Zeilinger, et al., 2013). Several were robustly replicated in multiple studies (Bojesen, et al., 2017; R. Philibert, et al., 2012), making AHRR the most popular biomarker of smoking habits (Maas, et al., 2019). Among the most studied sites, the CpG cg05575921 (annotation on the 450K array) located in an intronic enhancer of the gene displayed the highest methylation change (Joehanes, et al., 2016; Zeilinger, et al., 2013). Decrease in methylation (hypomethylation) of this site was associated with higher levels of smoking in several studies across European, African and Asian ancestries, both sexes and a range of ages (Dogan, et al., 2014; R. A. Philibert, et al., 2013; Shenker, et al., 2013a; Zeilinger, et al., 2013), as well as maternal smoking during pregnancy (Joubert, et al., 2012). It was suggested that AHRR (cg05575921) hypomethylation may also provide a clinically relevant prediction of future smoking-related morbidity and mortality (Bojesen, et al., 2017). The exact mechanism of how changes in AHRR DNA methylation regulate its expression is not known, but the decreased DNA methylation may mediate upregulation of the gene, reported in current smokers compared to non-smokers (Chatziioannou, et al., 2017; Parker, et al., 2017).

One of the most fascinating aspects of smoking-induced DNA methylation changes is their varied reversibility upon smoking cessation. While the great majority of DNA methylation signatures in the blood seem to revert to levels observed in non-smokers within 5 years of smoking cessation, some changes persist for more than 30 years (Guida, et al., 2015; Hoang, et al., 2024; Joehanes, et al., 2016). For example, hypomethylation of the AHRR (cg05575921)

reverts slowly after smoking cessation, with 22 years needed to reach the non-smoking levels (Ambatipudi, et al., 2016). Across different EWAS studies, 4267 unique CpGs were significantly associated with smoking cessation [see (Fang, et al., 2023) for a recent systematic review of epigenetic biomarkers of smoking cessation]. Examples of genes linked to CpG sites with persisting smoke-related changes include AHRR, TIAM2, PRRT1, F2RL3, GNG12, LRRN3 and APBA2 (Fang, et al., 2023; Hoang, et al., 2024; Joehanes, et al., 2016). The mechanisms behind the persistence of methylation changes are not fully clear, smoking-induced alterations in haematopoietic stem cells of the bone marrow were suggested as a potential explanation (Guida, et al., 2015).

Of note, it is not fully clear whether the "reversibility" of specific sites is caused by the loss of an acquired DNA methylation change or depletion of a specific cell type (and its methylation signature) from the investigated tissue or both. As most of the smoking EWAS were performed in mixed material (blood or tissue), there are two possible explanations for the observed results: 1) the smoke-related DNA methylation change is a true epigenetic remodelling event, where exposure to cigarette smoke leads to an acquired change in the DNA methylation pattern in the exposed cells or 2) the detected DNA methylation change is a reflection of the changed cellular composition of the studied tissue upon cigarette smoke exposure (e.g., reflecting the influx of inflammatory cells). Consequently, the reversibility of DNA methylation may be a true loss of an acquired methylation change or may reflect the depletion of a cellular component with characteristic methylation at a given locus (e.g., less inflammatory cells present in blood or lung tissue upon smoking cessation). Hence, adjusting for changes in cell-type composition in complex tissues is critical when analysing and interpreting findings from EWAS, as DNA methylation change in bulk tissues often manifests as an inflated signal due to

a shift in cell-type proportions between cases and controls (Houseman, et al., 2015; Teschendorff & Zheng, 2017).

Analysis of primary bronchial epithelial cells from never, current, and former smokers revealed that most tobacco smoke-driven gene expression changes were rapidly reversible (Beane, et al., 2007). Interestingly, a subset of genes displayed only slow or even persistent changes, mirroring the varied reversibility patterns in DNA methylation observed in earlier population studies. *In vitro* models of cigarette smoke exposure may help provide mechanistic insights into the kinetics and mechanisms of the reversibility of smoke-related epigenetic and transcriptional changes.

The key limitation of the EWAS studies is that they provide only correlative data, and do not allow the establishment of the causal link between smoke exposure, altered DNA methylation and changed gene expression. For this, functional studies are needed. In addition, longitudinal data quantitatively measuring DNA methylation after smoking cessation over time are still very limited, as most studies performed to date are cross-cohort studies associating DNA methylation with time since quitting. Carefully designed longitudinal studies are needed to validate these results and uncover the details of the different reversal kinetics among smoking-associated sites.

7.2. Remodelling of DNA methylation in chronic obstructive pulmonary disease (COPD).

Chronic obstructive pulmonary disease (COPD) is a prevalent chronic lung disease, affecting more than 250 million people worldwide, caused by exposure to cigarette smoke and other toxic gases. It is characterised by a progressive airflow obstruction caused by the remodelling of small airways and destruction of the lung parenchyma, known as emphysema (Barnes, et al., 2015; GOLD, 2023). Numerous genetic loci have been associated with COPD and lung

function (Cho, et al., 2014; Hancock, et al., 2010; Heinbockel, et al., 2018; B. D. Hobbs, et al., 2017; Sakornsakolpat, et al., 2019; Wain, et al., 2015; Wyss, et al., 2018), however, they explain only a small fraction of COPD risk. Current studies estimate that the genetic component of COPD is <3% (for monogenic risk) and combined polygenic risk scores may explain up to 38% of COPD susceptibility (Ragland, et al., 2019). Notably, the environmental cause of COPD (cigarette smoke) and the disease phenotypes, including aberrant cell differentiation (e.g., Goblet cell metaplasia) indicate the involvement of epigenetic mechanisms in COPD development. Indeed, numerous studies provided strong evidence for the association of dysregulated DNA methylation and COPD (**Table 2**) in the blood (Bermingham, et al., 2019; Busch, et al., 2016; Carmona, et al., 2018; Qiu, et al., 2012), sputum (Sood, et al., 2010), oral mucosa (Wan, et al., 2015), lung tissue (Morrow, et al., 2016; Sood, et al., 2010; Sundar, et al., 2017; Yoo, et al., 2015), bronchial brushings (Vucic, et al., 2014) and isolated lung fibroblasts (Clifford, et al., 2018; Schwartz, et al., 2023). Importantly, DNA methylation changes were associated with altered expression of genes and pathways important to COPD pathology, indicating their potential role in driving aberrant gene expression programs in COPD cells (Prada, et al., 2023; Schwartz, et al., 2023). In addition, a recent study suggested that DNA methylation changes may originate in early life (Kachroo, et al., 2020), further emphasising the potential role of DNA methylation in conferring long-term risk for chronic lung disease development.

Most studies profiled DNA methylation using complex material with mixed cell populations and thus could not identify the cellular origin contributing to the observed DNA methylation and expression changes in COPD. Hence, the precise epigenetic changes in COPD driving cell populations and their contribution to altered transcriptional patterns in COPD are still not well understood. Only three studies to date investigated DNA methylation changes in COPD

patients in isolated lung cells (Clifford, et al., 2018; Prada, et al., 2023; Schwartz, et al., 2023). Using Illumina 450K BeadChip Array, Clifford et al. identified 887 and 44 differentially methylated regions in parenchymal and airway fibroblasts of COPD patients, respectively (Clifford, et al., 2018). Our group has pioneered the use of high-resolution profiling to examine the DNA methylation landscape in COPD in purified cells. Using whole genome bisulfite sequencing, a genome-wide approach, we identified widespread DNA methylation changes in primary lung fibroblasts (Schwartz, et al., 2023) and alveolar type 2 cells (Prada, et al., 2023) isolated from a small cohort of COPD patients. DNA methylation changes were associated with dysregulated expression of key pathways involved in COPD pathology, like proliferation, inflammation, transforming growth factor β (TGFβ) and Wnt signalling, indicating that they may contribute to disease development. Using epigenetic editing, we provided the first evidence of the functional role of aberrant DNA methylation in driving interferon signalling in COPD (Prada, et al., 2023), highlighting the potential of cell-type resolved epigenetic profiling in identifying epigenetic drivers of disease phenotypes.

Little is known about the correlation of DNA methylation with disease severity, trajectories, or progression. DNA methylation changes in 13 genes have been identified in the lung tissue of COPD GOLD I and II patients compared to non-smoker controls (Casas-Recasens, et al., 2021). WGBS data from our group revealed that genome-wide DNA methylation changes are present in lung fibroblasts from COPD GOLD I patients (mild COPD) compared to controls with matched smoking status and history (Schwartz, et al., 2023), demonstrating that epigenetic changes occur already in mild disease. Some of these changes “progressed” with disease severity. However, another study found distinct DNA methylation signatures between mild and severe COPD patients, indicating a possible non-linear relation between DNA methylation and disease development (Casas-Recasens, et al., 2021).

These initial studies demonstrate that DNA methylation may provide a sensitive biomarker for COPD detection and patient stratification. Carefully designed longitudinal studies in larger cohorts are needed to validate these results obtained with the small number of samples and uncover the details of the epigenetic changes associated with different COPD subtypes and disease trajectories. The combination of epigenetic profiling with other NGS-based omics approaches (e.g. proteomics, metabolomics, transcriptomics, radiomics) holds promise for the Identification of COPD subtypes driven by common pathological mechanisms (endotypes) that are key to the development of curative therapies for COPD (Brian D. Hobbs, et al., 2023; Olvera, et al., 2024; Polverino & Kalhan, 2023) .

7.3. Remodelling of DNA methylation patterns in pulmonary fibrosis.

Lung fibrosis is caused by an aberrant lung tissue repair process characterised by excessive deposition of extracellular matrix and proliferation of fibroblasts (Lederer & Martinez, 2018). Among fibrotic lung disease, Idiopathic pulmonary disease (IPF) is an incurable lung disease, affecting 5 million people worldwide, with increasing prevalence and healthcare burden (Martinez, et al., 2017). It is a devastating disease leading to irreversible scarring of the lung, progressing breathing difficulties, and ultimately death 3-5 years after diagnosis (Lederer & Martinez, 2018; Raghu, et al., 2018). Despite the discovery of genes predisposing to IPF (Allen, et al., 2017), the molecular mechanisms involved in the initiation, development, and progression of IPF are unknown. Considering that most known risk factors for IPF, including age, cigarette smoke exposure, and male sex are associated with DNA methylation changes, DNA methylation is likely to play a role in IPF development. Changes in DNA methylation have been identified in lung tissue (Rabinovich, et al., 2012; Sanders, et al., 2012; Yang, et al., 2014) and fibroblasts (Huang, et al., 2014; Lee, et al., 2019) from IPF patients, providing the first

evidence of dysregulated epigenetic signalling in IPF. Yet, comprehensive profiling of epigenetic dysregulation in IPF has not been carried out.

Previous gene-specific studies using cellular models or lung tissues of IPF patients identified altered DNA methylation and expression of several genes important to IPF pathology. These include fibrogenic or anti-fibrotic genes, like Thy-1 antigen (THY1), Prostaglandin E receptor 2 (PTGER2), Caveolin 1 (CAV1), α-smooth muscle actin (α-SMA), Phosphatase and tensin homolog (PTEN), chemokine IP-10 and cyclin-dependent kinase 4 inhibitor B (CDKN2B), where a change of DNA methylation was associated with altered gene expression (Huang, et al., 2010; Sanders, et al., 2017; Sanders, et al., 2008; Scruggs, et al., 2018).

Several studies investigated DNA methylation changes using more comprehensive approaches (**Table 3**). Using low-coverage arrays (focusing on promoters), three pioneering studies provided evidence of widespread dysregulation of DNA methylation in IPF lung tissue and its association with gene expression changes (Rabinovich, et al., 2012; Sanders, et al., 2012; Yang, et al., 2014). The most comprehensive of the three examined 4.6 million CpG sites and identified 2,130 significant DMRs, 870 of which were associated with differentially expressed genes. (Yang, et al., 2014). DNA methylation changes were predominantly located in gene bodies and CpG island shores, with only 10% located in gene promoters (and outside CpG islands). IPF methylomes also did not exhibit hypomethylation of retrotransposons (Rabinovich, et al., 2012), suggesting that methylome features in IPF may differ from typical changes characteristic of cancer cells.

These studies profiled whole lung tissue and hence did not account for epigenetic differences in individual cell types. This was addressed by subsequent studies that identified alterations in DNA methylation in IPF using purified lung cells, including parenchymal fibroblasts (Huang,

et al., 2014; Lee, et al., 2019) and alveolar macrophages (AM) (McErlean, et al., 2021). The most recent one profiled AM isolated from bronchioalveolar lavage (BAL) of IPF patients and controls using Illumina EPIC array (McErlean, et al., 2021). The authors identified epigenetic heterogeneity as a key feature of AM in IPF, mirroring transcriptional heterogeneity of AM identified in single-cell RNA sequencing studies and confirming earlier results from IPF fibroblasts (Huang, et al., 2014). One challenge of DNA methylation studies in complex diseases like IPF is their heterogeneous nature, as different stages of the disease might have different epigenetic landscapes and even regions with active disease may differ from the areas with established fibrosis within the same lung. Indeed, hypermethylation of the THY1 promoter was observed in fibroblasts within active fibroblastic foci, but not in dense, fibrotic areas, indicating the temporal and spatial regulation of DNA methylation in IPF (Sanders, et al., 2008). In addition, rejected donor lungs are often used as controls for IPF studies as obtaining an age-matched normal cohort is challenging. As these tend to be younger; some of the identified changes may be due to ageing as the DNA methylation landscape changes with age and the chronological age can be quantified by DNA methylation patterns (Horvath & Raj, 2018).

Enrichment of DNA methylation changes outside promoters and CpG islands makes inferring their functional relevance more difficult, as complex interactions exist between epigenetic mechanisms, chromatin structure, nuclear architecture, and gene expression. Functional studies are urgently needed to evaluate the role of identified alterations in disease pathology and to understand the interaction between genetic predisposition and epigenetic regulation in IPF.

Overall, DNA methylation profiling studies provide strong evidence that widespread alterations in DNA methylation are associated with the dysregulation of genes important in the pathogenesis of IPF in the lung.

7.4. Alterations of DNA methylation patterns in asthma.

Asthma is the most common chronic lung disease, affecting around 300 million people worldwide (Porsbjerg, et al., 2023). It is characterised by variable respiratory symptoms and reversible airway obstruction. In contrast to COPD and IPF, for which age is a significant risk factor, asthma can affect people of all ages. There is strong evidence indicating that the aetiology of asthma, as well as its clinical course results from complex interactions between host genotype and environmental exposures (Bonnelykke & Ober, 2016; Melen, et al., 2022). Early-life exposures, including intrauterine exposure to cigarette smoke, viral infections or maternal diet are known risk factors, highlighting the important role of environmental exposures in asthma pathology. Among chronic lung diseases, asthma is the one with the best-established association with epigenetic dysregulation, as altered DNA methylation patterns were associated with allergy, and atopy in both children and adults.

A few selected examples illustrating the contribution of dysregulated DNA methylation to asthma are discussed below (**Table 4**), but the reader is encouraged to refer to (Edris, et al., 2019; Legaki, et al., 2022; Sheikhpour, et al., 2021) for recent systematic reviews of asthma-associated DNA methylation changes in a variety of relevant clinical samples. Several genes and CpGs with altered DNA methylation have been suggested as potential asthma biomarkers. Among the most stably replicated were: interleukin 5 receptor subunit alpha (IL5RA), eosinophil peroxidase (EPX), SMAD3 family member 3 (SMAD3)and RUNX family transcription

factor 3 (RUNX3) (Cardenas, et al., 2019), key genes involved in T cell maturation, Th2 immunity and asthma pathology.

Two large-scale meta-analyses investigated the association of DNA methylation changes in blood to childhood asthma (Reese, et al., 2019; C. J. Xu, et al., 2018). Using data from more than 5000 children from six European cohorts, Xu and colleagues (2018) reported reduced DNA methylation at 14 specific CpG sites, strongly associated with childhood asthma (C. J. Xu, et al., 2018). The effects were strongly driven by lower DNA methylation within eosinophils, demonstrating the key role of this cell population as an epigenetic contributor to childhood asthma. In turn, Rheese et al. (2019) identified 9 CpGs and 35 regions differentially methylated in newborn blood that may be potential biomarkers of risk for asthma development (Reese, et al., 2019).

DNA methylation profiling of airway epithelial cells in adult asthmatics was used to identify epigenetic signatures of distinct disease endotypes, reflecting key components of asthma pathogenesis: airway remodelling, eosinophilia and nitride oxide (NO) response (Nicodemus-Johnson, et al., 2016). The authors concluded that DNA methylation profiles constitute a more stable disease biomarker than transcriptional signatures (Nicodemus-Johnson, et al., 2016). A recent study in whole blood characterised the complete genome-wide DNA-methylation patterns associated with childhood asthma using WGBS (Thurmann, et al., 2023). It reported a prominent loss of DNA methylation (hypomethylation) affecting predominantly enhancer elements, associated with key immune genes involved in asthma pathology, demonstrating the value of unbiased, high-resolution DNA methylation profiling.

The nasal epithelium is a promising tissue for studying DNA methylation changes related to asthma, as demonstrated in several genome-wide studies (**Table 4**) (Cardenas, et al., 2019;

740 Forno, et al., 2019; X. Zhang, et al., 2018). Interestingly, many asthma-associated CpG sites
741 discovered in whole blood replicated in nasal epithelial cells, probably reflecting common
742 inflammatory processes (Cardenas, et al., 2019). While genes associated with DNA
743 methylation changes in blood samples were involved inflammatory mediators, these
744 identified in nasal cells also included many proteins of extracellular matrix and membrane
745 proteins. Nasal epithelia are easier to access than blood samples and might be a better proxy
746 for pathological changes in airway cells [reviewed in (Solazzo, et al., 2020)]. However, the nasal
747 epithelium is still a complex material containing several cell types, hence some of the obtained
748 profiles may reflect differences in cell type composition of diseased and not diseased samples.
749 Indeed, after cell-type adjustment, many effects were markedly reduced (Cardenas, et al.,
750 2019).

751 **8. Occupational exposures modify DNA methylation and the risk of diseases**

752 The association of DNA methylation changes with adverse health effects caused by exposure
753 to environmental and occupational toxicants (e.g., particulate matter, cigarette smoke, diesel
754 exhaust fumes, ozone, silica, coal dust, asbestos and a variety of other toxic chemicals and
755 metals, among others) is well documented in experimental and epidemiological studies [see
756 (Martin & Fry, 2018; Nwanaji-Enwerem & Colicino, 2020) for comprehensive general reviews].
757 Animal models provide emerging evidence that occupational exposures may affect not only
758 the lungs but also other organs, including the brain and are associated with epigenetic
759 dysregulation and neurodegeneration (Shoeb, et al., 2020). Notably, some environmentally
760 induced epigenetic changes may even be heritable across generations (Anway, et al., 2005;
761 Greger, et al., 1989), potentially contributing to familial aggregation of chronic diseases.
762 Hence, it is not surprising that occupational exposures not only increase the risk of developing

763 chronic lung diseases, including COPD, IPF and asthma but also lead to disease exacerbations,
764 as documented in population association studies as well as animal models [see (Dao &
765 Bernstein, 2018; Gandhi, et al., 2024; Murgia & Gambelunghe, 2022; Walters, 2020) for
766 reviews].

767 Due to the strong association between epigenetic changes and exposures, DNA methylation
768 has the potential to serve as a measurable biomarker of exposure to occupational or toxic
769 agents and a possible mediator of exposure effects, leading to the development of the concept
770 of “toxicomethylomics“ (Szyf, 2011) [or “toxicoepigenetics” in general]. However, several
771 limitations need to be considered before alterations in DNA methylation can be used as
772 biomarkers for environmental or occupational exposures (Svoboda, et al., 2022).

773 One of the main challenges in environmental epigenetic toxicology is the complex interaction
774 between various environmental factors, diet and hormone signalling, which are all associated
775 with DNA methylation changes. For example, a high-fat diet may change the susceptibility of
776 the lung response in individuals exposed to specific occupational exposures, as documented
777 in animal models (Antonini, et al., 2019). Another important consideration is the time and
778 duration of exposure. Because changes in DNA methylation can persist for years, as
779 highlighted above for cigarette smoke, it is difficult to establish a causal relationship between
780 the initial exposure and the development of a given disease, as there could be a substantial
781 lag between the two. In addition, due to the unavailability of the disease-relevant tissue, most
782 methylation-based biomarkers of environmental/occupational exposures were derived using
783 blood, often with limited representation of different ethnicities in the studied populations.
784 Therefore, there may be limited translatability of blood biomarkers to other populations or to
785 target organs.

The establishment of prospective and ethnically diverse cohorts, with longitudinal analysis of different sample types is required to advance epigenetic toxicology research. There is an urgent need for the inclusion of detailed nutritional and environmental assessments in the studies to better characterise the samples and account for complex exposures (the exposome) and interactions. Sex is a well-known factor in the development of chronic lung diseases, hence sex-based differences in DNA methylation should also be investigated because they may influence the response to environmental exposure (Huen, et al., 2014) and mediate disease risk.

A key research priority is to determine whether environmentally induced epigenetic alterations may have a causative effect on disease development. For this, the functional implication of the identified changes needs to be investigated in experimental models using innovative technologies, like epigenetic editing described above. Mechanistic understanding of how environmental/occupational exposures impact the epigenome to cause/exacerbate chronic lung diseases may enable the development of better screening assays and inform policies for the safety assessment of agents and occupational hazards, as well as open new avenues for non-invasive biomonitoring and prevention strategies.

9. Evolving concepts in DNA methylation

Genome-wide DNA profiling studies provided strong evidence of dysregulated DNA methylation in response to smoke exposure and chronic lung diseases, including COPD, IPF, asthma and others. Interestingly, they found that most differentially methylated regions/sites are located outside gene promoters and are not enriched in CpG islands, as previously reported in cancer (Jones, 2012). Instead, an enrichment in gene bodies, CpG islands shores, intergenic regions and distal enhancers have been observed (Hoang, et al., 2024; Prada, et al.,

2023; Rabinovich, et al., 2012; Schwartz, et al., 2023; Thurmann, et al., 2023). Hypomethylation of enhancer regions emerges as a key feature in COPD (Prada, et al., 2023; Schwartz, et al., 2023) and asthma (Thurmann, et al., 2023), indicating that DNA methylation at enhancers may play a pivotal role in the pathogenesis of chronic lung diseases.

DNA methylomes from chronic lung diseases also did not exhibit hypomethylation of retrotransposons typically observed in cancer cells (Rabinovich, et al., 2012). These observations highlight the differences in the DNA methylation landscape between cancer and chronic lung diseases, indicating disease specificity despite shared risk factors. They also evidence that genome-wide methods (like WGBS) with more comprehensive coverage are urgently needed to characterise the DNA methylation landscape of chronic lung diseases.

Another important observation came from the investigations of the correlation between DNA methylation and gene expression. As DNA methylation is a repressive epigenetic signal, methylation changes are often expected to be inversely correlated with gene expression (hypomethylation with increased expression and hypermethylation with repression, respectively). Hence, a positive correlation between gene expression and DNA methylation was previously considered contradictory. With the recent high-resolution mapping of DNA methylation and gene expression across multiple tissues and cells, it has become increasingly clear that DNA methylation has complex and context-dependent roles. The canonical negative correlation is mostly observed in gene promoters (Weber, et al., 2007), although examples of positive correlation were also observed (Prada, et al., 2023; Spainhour, et al., 2019). The mechanism can be explained by the recruitment and binding of transcriptional repressors that show higher affinity to methylated DNA (so-called methyl Plus TF) or by DNA methylation inhibiting the binding of activating complexes [reviewed in (Greenberg & Bourc'his, 2019)]. In

turn, DNA methylation in gene bodies is associated with high gene expression (Lister, et al., 2009; Varley, et al., 2013). There, it prevents spurious transcription activation from alternative promoters, contributing to higher gene expression efficiency. The varied role of DNA methylation across different genomic regions has important implications for understanding the role of altered DNA methylation in disease, as its effects will be context and location dependent.

Finally, with the availability of genome-wide methylome across tissues and disease states, it became clear that methylome studies from complex tissues need to be interpreted cautiously, as cellular heterogeneity is an important confounder in DNA methylation (also in gene expression studies) (Houseman, et al., 2015; Teschendorff & Zheng, 2017). For example, whole blood comprises at least 7–8 main cell types, while human lung tissue contains 58 different cell types (Adams, et al., 2023), each with a unique DNA methylation pattern. Hence, cellular composition variations between tested sample groups (e.g., due to disease process) can confound DNA methylation analyses. Indeed, it has been demonstrated that the observed DNA methylation effects (e.g. positive associations in EWAS) can often result from an inflated signal due to a shift in cell type proportions between the cases and controls (Jaffe & Irizarry, 2014). Similarly, a recent study provided strong evidence that differences in the proportion of naïve and differentiated subsets of CD4+ and CD8+ T cells is a key factor contributing to DNA methylation variation in the blood that may mediate associations between DNA methylation and environmental exposures (Bergstedt, et al., 2022). Therefore, validation in purified cell populations or deconvolution of DNA methylation data to adjust for cell type proportions should be a prerequisite for epigenetic association analysis. A variety of deconvolution strategies and statistical methods have been developed and systematically evaluated (Jeong, et al., 2022; Song & Kuan, 2022) [refer to (Teschendorff & Zheng, 2017; Titus, et al., 2017) for

a critical review of available approaches]. In general, deconvolution can be performed reference-free, using reference methylomes from the cell types of interest (Teschendorff & Zheng, 2017) or more recently using single-cell RNA seq data (Teschendorff, et al., 2020; Zhu, et al., 2022). To enable further progress, we need to generate high-resolution genome-wide reference epigenetic profiles of healthy and diseased lung cells isolated from well-clinically characterised biomaterials. The recent development of novel workflows based on tissue cryopreservation could enable the biobanking of viable lung tissue for future cell isolation and profiling (Llamazares-Prada, et al., 2021; Pohl, et al., 2023). In addition, several novel technologies have recently been developed to map DNA methylation and other epigenetic modifications with single-cell and even spatial resolution (Ahn, et al., 2021; Baysoy, et al., 2023). Although not mature yet and applied to selected contexts only, further development of such technologies will enable the investigation of epigenetic modifications with single-cell and spatial resolution, as is already the case for transcriptomic research.

10. Mechanism of epigenetic changes in chronic lung diseases.

Despite substantial progress in mapping epigenetic alterations in chronic lung diseases, the molecular mechanisms leading to their establishment are largely unknown. Due to the vital role of DNMTs and TET enzymes, their recruitment and activity in cells need to be tightly controlled. Several models have been proposed to explain how specific DNA methylation patterns are established and edited in cells [reviewed in (Jurkowska & Jurkowski, 2019)]. They highlight the key role of protein partners and chromatin modifications as well as the contribution of the inherent properties of the enzymes themselves.

Changes in DNA methylation at specific genomic regions can result from the altered expression, localisation or activity of the enzymes that set (DNA methyltransferases) or

remove (TET enzymes) the methylation marks. In addition, DNA methylation cross-talks to other epigenetic layers, therefore, changes in histone modifications and chromatin structure will affect DNA methylation and vice versa. Numerous examples of all these mechanisms have been documented across different tissues and disease states, but little is known about them in the context of chronic lung diseases.

Overexpression of DNA methyltransferases (or their splicing isoforms) has been observed in multiple cancers, including lung cancer. Altered expression of DNMT1, DNMT3A, DNMT3B and MeCP2 (methylated DNA reader) were reported in IPF lung tissue and fibrotic lungs of bleomycin-treated mice (Sanders, et al., 2012; Wei, et al., 2022) and in response to cigarette smoke condensate in cells (F. Liu, et al., 2010). These selected examples indicate that transcriptional dysregulation of the writers, readers and erasers of DNA methylation occurs also in chronic lung diseases, providing a first mechanistic link for the observed global alterations in the DNA methylation landscape reported in profiling studies. However, the molecular mechanism of how altered expression of epigenetic factors is established by environmental exposures remains unknown.

10.1. How are DNA methyltransferases and TET recruited?

In addition, DNMTs and TETs get recruited via interactions with specific chromatin modifications, transcription factors or other protein partners [reviewed in (Jurkowska & Jeltsch, 2022a)].

Chromatin modifications emerged as one of the key mechanisms for recruiting DNA methyltransferases to specific genomic regions. All human MTases possess specific reading domains in their N-terminal domains that can directly recognise and bind specific histone modifications, bringing the MTase activity to specific genomic regions. For example, DNMT3A

and DNMT3B methyltransferases interact via their ATRX-DNMT3-DNMT3L (ADD) domains with histone H3 tails unmodified at Lysine 4 (Otani, et al., 2009; Y. Zhang, et al., 2010). Notably, H3 binding allosterically activates the enzyme, stimulating methylation of the neighbouring DNA (Guo, et al., 2015). This is a beautiful example of the cooperation between different epigenetic signals, whereby one epigenetic modification can lead to the establishment of another mark in the same region. Conversely, H3K4me4, which marks active gene promoters, blocks the binding of DNMTs, preventing DNA methylation of these active regions and keeping them open to transcriptional machinery. These are just two selected examples illustrating the important role of chromatin modifications in the recruitment and regulation of DNA methyltransferase and demethylase machinery, for more details about the chromatin interaction of DNMTs, the reader is referred to a recent review (Jurkowska & Jeltsch, 2022a).

Numerous protein partners that can recruit epigenetic enzymes to specific genomic regions and directly modulate the activity have been identified (Jurkowska & Jeltsch, 2022a). For example, MeCP2, a methyl-binding protein that can recognise methylated CpG sites, directly interacts with and inhibits the activity of DNMT3A methyltransferase via an allosteric mechanism (Rajavelu, et al., 2018). Similarly, UHRF1, which is required for efficient maintenance of DNA methylation during DNA replication, was shown to both recruit DNMT1 maintenance MTase to the replicating chromatin and at the same time to stimulate the activity of the enzyme (Bashtrykov, et al., 2014), providing another example of the multifaceted regulatory mechanism controlling the activity and localisation of these important enzymes.

Finally, environmental exposures may influence the activity of epigenetic enzymes by regulating the availability of their required cofactors. For example, TET demethylases and the histone lysine demethylating (Jumanji) proteins are oxygen-dependent enzymes. Thus, oxygen

levels can directly influence their catalytic activity, which is particularly relevant in chronic lung diseases often associated with hypoxia. Reduced TET activity due to hypoxia led to DNA hypermethylation of gene promoters in cancer (Thienpont, et al., 2016). Hence, hypoxia emerges as one of the regulators of DNA methylation patterns in disease. Similarly, levels of methyl cofactors for DNA and histone methylation (SAM) can be regulated by the availability of folate, coupling metabolism and diet to epigenetic regulation.

10.2. How do altered DNA methylation patterns translate into biological effects?

How altered DNA methylation patterns translate into biological effects in lung cells is currently unclear. DNA methylation in regulatory regions (e.g. promoters, enhancers, insulators) can modulate the binding of transcriptional activators and repressors to DNA (Stadler, et al., 2011). Indeed, DNA methylation within the binding sequence of a transcription factor (TF) can directly prevent their binding (these are so-called Methyl minus TF), leading to the loss of their function at the methylated regions. Conversely, methylation of DNA may enhance the binding of selected TFs (Methyl plus TF), providing a potential gain of function mechanism for downstream signalling (Yin, et al., 2017). In addition, DNA methylation (together with repressing histone modifications) usually leads to chromatin compaction, making DNA inaccessible to transcriptional machinery. An elegant example of the prominent role of DNA methylation in contributing to lung disease risk was recently provided by Helling and colleagues (Helling, et al., 2017), who investigated the molecular mechanism behind the mucin 5B (MUC5B) promoter variant (rs35705950), the largest genetic risk factor for IPF development. They revealed that the variant is associated with regional changes in DNA methylation of an enhancer element and MUC5B expression, which may be mediated by the transcription factor FOXA2 (Helling, et al., 2017).

Of note, whole-genome DNA methylation or chromatin accessibility profiling (e.g. by assay for transposase-accessible chromatin with sequencing (ATAC-seq)) allows for inferring transcriptional regulators potentially mediating (or responding to) the epigenetic alterations in regulatory regions (Stadler, et al., 2011). Our recent WGBS methylation study in fibroblasts reported a significant enrichment of binding sites for TCF21 and FOSL2/FRA2 transcription factors in the differentially methylated regions overlapping with strong enhancers in COPD lung fibroblasts, identifying them as potential disease regulators and mediators of epigenetic changes in cells (Schwartz, et al., 2023). However, follow-up studies directly footprinting the binding of TF in healthy and diseased cells (e.g. by chromatin immunoprecipitation with sequencing (ChIP-seq)) combined with their genetic manipulation (e.g knockdown or overexpression) are needed to unravel the mechanisms of epigenetic regulation in response to DNA methylation alteration.

11. Epigenetic mechanisms as potential therapeutic targets – promises and challenges.

Unlike disease-associated genetic mutations, epigenetic alterations are reversible and hence constitute attractive intervention targets. Epigenetic-based therapies aim to restore the normal epigenetic state and reverse aberrantly activated or silenced genes. The main promise of this approach is that it targets the epigenetic regulatory layer itself, hence correcting gene expression dysregulation at its source without making changes to the underlying genetic sequence. With the fine mapping of specific epigenetic alterations in human diseases and the advancement in mechanistic understanding of epigenetic enzymes, came the idea that epigenetic factors may provide useful targets for clinical applications. All three types of epigenetic proteins—writers, readers, and erasers—are in principle druggable and can be targeted through small-molecule inhibitors.

Cancer is the best-known example of a disease with a demonstrated benefit of epigenetic modulation as a therapeutic approach. Several small-molecule inhibitors targeting histone- and DNA-modifying enzymes (e.g. DNA methyltransferases, histone deacetylases and histone methyltransferases) have been developed. Eight have been approved by the Food and Drug Administration (FDA). Seven are currently used in clinics, five for the treatment of hematologic malignancies and two for solid tumors (Feehley, et al., 2023) (**Table 5**). Many others are widespread in clinical trials for malignancies and other diseases (Feehley, et al., 2023; Ganesan, et al., 2019; X. Yu, et al., 2024).

11.1. Targeting DNA methyltransferases

DNA methyltransferases can be inhibited by targeting different parts of the catalytic pocket (e.g. DNA binding site, SAM cofactor binding sites or both) or by targeting the allosteric regulatory regions outside the enzymatic domain. The field of epigenetic-based therapy was pioneered by Peter Jones with the development of 5-azacytidine and 5-aza-deoxycytidine (decitabine) (Jones & Taylor, 1980), the first two DNA methyltransferase inhibitors (DNMTi) approved for the treatment of haematological cancers. 5-azacytidine (and its analogues) get incorporated into the DNA and form an irreversible covalent complex with DNA methyltransferases, often referred to as a suicidal complex. This leads to the degradation of DNMTs in cells, loss of DNA methylation and re-expression of tumour suppressor genes previously silenced by DNA methylation. However, despite their efficacy, both drugs have poor stability, low bioavailability, lack selectivity and are associated with relatively high toxicity, limiting their clinical utility (Ganesan, et al., 2019; Ma & Ge, 2021). Later, other modalities, for example, the methyl donor cofactor (SAM) analogues were also introduced to inhibit DNA methyltransferases without inducing DNA damage. Today, many derivatives of these initial compounds with improved stability and pharmacokinetics have been developed (e.g.

995 Guadecitabine, SGI-110) and are tested in clinical trials but are not yet in clinical practice
996 (Feehley, et al., 2023; Ganesan, et al., 2019; X. Yu, et al., 2024). Further development led to
997 the idea of combination treatments, where DNMT inhibitors could be combined with
998 additional epigenetic or antitumor drugs, resulting in higher efficacy (Hu, et al., 2021).

999 **11.2. Targeting histone-modifying enzymes**

1000 Another class of epigenetic drugs used in oncology targets histone-modifying enzymes, with
1001 the most prominent example of histone deacetylase inhibitors (HDACis). HDACi block histone
1002 deacetylases (HDAC), enzymes that remove acetyl groups from lysine residues on histones,
1003 leading to chromatin compaction, and gene silencing. HDACis counteract the abnormal
1004 acetylation exhibited by cancer cells, leading to gene reactivation (Falkenberg & Johnstone,
1005 2014). One of the mechanisms of HDACis is the activation of apoptosis pathways that impede
1006 the growth and survival of tumour cells (Matthews, et al., 2012). To date, several generations
1007 of HDAC inhibitors with various chemistries have been developed, most of which act by
1008 binding the zinc cation in the HDAC active site. Five have been approved by the FDA
1009 (vorinostat, belinostat, romidepsin, tucidinostat and panobinostat) for the treatment of
1010 various indications, mostly haematological neoplasms (Bondarev, et al., 2021). Among them,
1011 vorinostat and romidepsin were the first approved epigenetic drugs that targeted
1012 modifications of histones (Ganesan, et al., 2019).

1013 Interestingly, the utility of epigenetic therapy is also actively investigated in the context of viral
1014 infections aiming to reactivate epigenetically silenced latent viruses (e.g., human
1015 immunodeficiency virus (HIV1) or human cytomegalovirus (HCMV)) to enable their
1016 eradication by immune cells or antiviral therapy. Several Phase I/II clinical trials have been

conducted to evaluate the effect of HDACi (and other epigenetic drugs) in the reactivation of HIV1 viral latency (Nehme, et al., 2019).

Like DNMTi, the pharmacokinetic profile of HDACi is not optimal, and they display off-target effects due to non-selective metal binding, hence the search for more specific next-generation HDACi continues. As epigenetic processes are interconnected, combining several epi-drugs might be a promising approach to epigenetic therapy. A novel strategy addressing the low selectivity and efficacy of epigenetic inhibitors is the use of bifunctional compounds. This approach involves the design of inhibitors with dual specificity which can be achieved by linking two (or more) active moieties in a single drug.

11.3. Targeting epigenetic readers

Epigenetic writers and erasers are not the only classes of druggable epigenetic factors. The third group of potential targets contains the epigenetic reading domains, which bind epigenetic modifications and translate them into biological signals. Unlike inhibiting the catalytic activity of epigenetic enzymes, inhibitors of epigenetic readers (e.g. histone binding modules) typically disrupt protein-protein interactions. Two seminal studies in 2010, provided an elegant proof of concept for this approach with the development of bromodomain inhibitors JQ1 (Filippakopoulos, et al., 2010) and the synthetic compound I-BET (Nicodeme, et al., 2010). Bromodomains are widespread epigenetic reading modules present in more than 40 human chromatin proteins that specifically recognise acetylated lysines in histones (and other proteins). I-BET showed potent anti-inflammatory effects by inhibiting the binding of bromodomain-containing Bromodomain and Extra Terminal domain (BET) proteins to acetylated histones, thereby disrupting chromatin complexes essential for the expression of key inflammatory genes (Nicodeme, et al., 2010).

In addition, epigenetic drugs have been successfully tested in combination with chemotherapy, radiotherapy, hormone therapy, anti-angiogenic therapy, or immunotherapy [reviewed in (Morel, et al., 2020)], indicating that beyond their potential as monotherapies, they could have significant synergistic roles with other anticancer therapies.

In the context of lung diseases, lung cancer is not the only therapeutic area that could benefit from such epigenetic interventions, as many chronic lung diseases, including COPD, asthma, and lung fibrosis, have been associated with epigenetic and transcriptional dysregulation (as described above). Several studies provided initial evidence supporting the potential benefit of targeting epigenetic processes in lung diseases. A few examples are discussed below.

Treatment with a combination of DNA methyltransferase inhibitor and trichostatin A, an HDAC inhibitor, reduced the mortality rate, lung inflammation and lung injury in a lipopolysaccharide (LPS)-induced mouse model of acute lung injury (ALI) (Thangavel, et al., 2014). Targeting DNMT1/DNMT3a and the peroxisome proliferator-activated receptor-γ (PPAR-γ) axis with 5-aza led to demethylation of the PPAR-γ promoter, restored PPAR-γ expression, and alleviated lung fibrosis (Wei, et al., 2022).

Although epigenetic therapy is a promising strategy for the treatment of cancers and other lung diseases in the future, important challenges remain to be solved. Despite their use in clinics, epigenetic drugs demonstrate poor pharmacokinetics and low tolerability, mainly due to their lack of genomic and cell-type specificity. Hence, novel approaches with increased precision are needed to advance the epigenetic therapy field.

11.4. Epigenetic editing

Multiple EWAS studies identified DNA methylation alterations associated with lung function, smoke exposure or chronic lung diseases, suggesting that epigenetic signalling may play a

pathogenic role in lung disease development. However, the key limitation of the EWAS studies (and most current epigenetic studies in basic research) is that they provide only correlative data, and do not allow the establishment of the causal link between environmental exposure (e.g. smoking), altered DNA methylation and the observed phenotype. Thus, currently, we do not know which epigenetic alterations are the cause and which are the consequence of the disease process. Therefore, it remains unclear whether epigenetic mechanisms can be targets for novel therapeutic approaches for lung regeneration. With the advent of designer epigenetic editing technologies, we may finally be able to address that challenge and provide direct evidence for the pathogenic role of DNA methylation in disease development.

Epigenetic editing is an exciting new technology to revert epigenetic modifications at a genomic region of interest [reviewed in (Jeltsch, et al., 2007; Jurkowski, et al., 2015; Nakamura, et al., 2021)]. It employs a programable DNA targeting domain fused to an epigenetic effector domain, which can be specifically targeted to a desired gene to change its epigenetic state, and consequently its expression (**Figure 7**). Upon delivery into target cells, the epigenetic programmer is recruited to the target region (e.g. gene promoter or enhancer) via sequence-specific interaction of the DNA-binding domain, enabling the co-delivery of the fused epigenetic domain (e.g., DNA methyltransferase, demethylase or transcriptional regulator). The epigenetic effector domain will then exert its enzymatic (or transcriptional) activity, leading to gene activation or silencing.

Several different DNA binding platforms have been engineered to enable locus-specific targeting of epigenetic domains. These include the modular zinc-finger (ZF) transcription factors (Wolfe, et al., 2000), transcription activator-like effectors (TALEs) (Boch, et al., 2009) and more recently the catalytically inactive Cas9 (dCas9) nuclease (Qi, et al., 2013) from the RNA-directed clustered regulatory interspaced palindromic repeats (CRISPR) system

[reviewed in (Jurkowski, et al., 2015)]. dCas9 is the most exciting addition to the epigenetic targeting toolbox due to the simplicity of its target design, as directing Cas9 to a specific genomic location only requires designing a guide RNA (gRNA) specific for the desired target (Hsu, et al., 2014).

A variety of epigenetic effector domains for efficient engineering of transcriptional states in living cells or organisms have been designed and validated [reviewed in (Lau & Suh, 2018; Nakamura, et al., 2021)]. These range from transcriptional activators (e.g. VP64 (Perez-Pinera, et al., 2013) or tripartite VPR (Chavez, et al., 2015)) or repressors (e.g. KRAB), through epigenetic enzymes or their isolated domains (e.g. DNA methyltransferases, demethylases or histone modifying enzymes) or even more complex fusions consisting of multiple functional domains joined by a linker (e.g. engineered Dnmt3a-Dnmt3L (Stepper, et al., 2017) or Dnmt3a-KRAB fusion proteins).

Targeting DNA methyltransferases (or demethylases) to specific genomic loci provides unique tools to investigate the causal role of DNA methylation of gene expression and explore the potential role of epigenetic editing for epigenetic therapy. The first proof of concept for this approach was provided by a pioneering study by Xu and Bestor in 1997 (G. L. Xu & Bestor, 1997). Since then, epigenetic editing has been successfully used to study program epigenetic states in different cells (Chavez, et al., 2015; X. S. Liu, et al., 2016; Saunderson, et al., 2017; Song, et al., 2017) and pre-clinical mouse models (Horii, et al., 2020; Liao, et al., 2017; Matharu, et al., 2019; Zhou, et al., 2018), offering exciting prospects for future therapeutic interventions, including lung diseases (D.-D. Wu, et al., 2018).

Epigenetic editing can also further our understanding of the interplay between genetic and epigenetic disease contributions. Most of the identified disease-associated genetic variations

lie outside protein-coding regions, therefore, they likely affect gene regulation by interplaying with epigenetic mechanisms.

While currently used mostly as a research tool to investigate the functional relevance of epigenetic alterations on gene expression and disease phenotypes, precision epigenetic editing approaches may not be that far from clinical application. Following the historic approval of the first CRISPR-based therapeutics for gene therapy, epigenetic editing is gaining momentum as well. Several epigenomic modulators are currently in preclinical development, with the first one already in clinical trials (**Table 6**). The appealing aspect of epigenetic regulation is that once the altered epigenetic signal is corrected, cellular epigenetic mechanisms will maintain the new state of the locus, such that it can be inherited through cell divisions. Hence, theoretically, epigenetic editing could be used as a single, "hit-and-run" intervention.

The FDA has recently granted an orphan drug designation to OTX-2002, a first-in-class "epigenomic controller", for the treatment of hepatocellular carcinoma (HCC) (**Table 6**). OTX-2002 is a messenger RNA (mRNA) molecule encoding ZF-DNMT and ZF-KRAB proteins, encapsulated in a lipid nanoparticle. It aims to epigenetically inactivate the expression of MYC oncogene, a master transcription factor regulating cell proliferation, differentiation, and apoptosis of cancer cells. OTX-2002 is currently in Phase I/II clinical trial (NCT05497453) that evaluates the safety, tolerability, and preliminary antitumor activity in patients with HCC (ClinicalTrials.gov, 2022; Rodriguez-Rivera, et al., 2023). Although the trial is ongoing and the results have not been published yet, Omega Therapeutics reported that treatment with OTX-2002 in the first eight patients resulted in intended epigenetic state change and c-MYC downregulation.

Another precision epigenomic programmer EPI-321, based on catalytically inactive dCas9, is being developed by Epic Bio. It recently obtained the FDA orphan drug designation as a treatment for facioscapulohumeral muscular dystrophy (FSHD), the most common form of adult muscular dystrophy. The company plans to start a Phase 1/2 clinical trial of EPI-321 to assess its safety, activity, and preliminary efficacy in people with FSHD in 2024. Several additional programs at Epic Bio are in the pipeline (https://epic-bio.com/). Excitingly for the lung field, one seeks to address alpha-1 antitrypsin deficiency, a genetic cause of COPD/emphysema. An epigenetic programmer (EPIC-341) is being developed to epigenetically suppress the endogenous mutated gene and replace it with an exogenous functional version of the gene.

12. Future directions

Despite the growing recognition of the contribution of DNA methylation (and other epigenetic mechanisms) to the pathology of lung diseases and substantial progress in mapping DNA methylation changes across a spectrum of lung conditions, many basic questions still await answers and future challenges emerge from our current knowledge. How is the aberrant DNA methylation (and other epigenetic modifications) established in response to environmental insults? How are DNA MTases and TET enzymes targeted and regulated in lung diseases? Which biological functions play the oxidized forms of 5-methylcytosine (5hmeC, 5fC and 5caC) in the pathology of lung diseases? Can we use epigenetic editing to repair disease-causing epigenetic states and target chronic lung diseases? These questions cannot be answered now as lung epigenetic research is still in its infancy, but they will inspire new generations of scientists to further explore epigenetic regulation in the context of chronic lung disease.

Acknowledgements:

The work in the author's laboratory is funded by the School of Biosciences (Cardiff University), the Academy of Medical Sciences Springboard Award [SBF007\100176] and the UKRI Future Leader Fellowship [MR/X032914/1] to RZJ.

Conflict of Interest Statement

The author has no conflicts of interest related to this publication.

Figure legend

Figure 1 Different layers of epigenetic modifications regulating chromatin structure.

Schematic diagram of the main epigenetic mechanisms, including DNA methylation (displayed as a pattern of methylated – red lollipops and unmethylated – white lollipops CpG sites), posttranslational modification (PTMs) of the histone tails, long non-coding RNA (lncRNA) and chromatin remodelling. Different epigenetic modifications recruit specific reader proteins that elicit biological responses. Methylated DNA (5mC) reader protein and histone PTM reader are shown as two examples.

Figure 2 Cytosine modifications present in human DNA

Schematic representation of the unmethylated cytosine (C), methylated cytosine (5mC), as well as its oxidised forms 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), generated by the TET enzymes. Parts of the figure were adapted from (Ravichandran, et al., 2018) with permission from the Royal Society of Chemistry.

Figure 3 DNA methylation cycle: establishment, maintenance and removal of DNA methylation

DNA methylation pattern is generated on unmethylated cytosines by DNMT3A and DNMT3B with the help of the stimulatory factor DNMT3L during the development and establishment of germ cells (blue). It is maintained during DNA replication by a maintenance enzyme DNMT1 (orange), which recognises hemimethylated DNA, consisting of a methylated parental strand (red lollipops) and an unmethylated daughter strand (white lollipops) and returns it to its

original pattern. DNA methylation can be lost by a passive mechanism when DNMT1 activity is absent. It can also be actively removed by subsequent oxidation of 5mC to 5hmC, 5fC and 5caC by TET enzymes, followed by the removal of 5fC and 5caC by TDG and base excision repair (BER) enzymes, leading to the restoration of unmethylated DNA.

Figure 4 The histone code

Selected examples of the most common posttranslational modifications of the N-terminal tails of core histones are displayed as green (acetylation), red (methylation) and yellow (phosphorylation) circles. Numerous additional modifications have been described, refer to the text for details.

Figure 5 Basic principles of epigenetic regulation of gene expression.

Active chromatin is open, decorated with activating epigenetic marks like acetylated histones, is usually devoid of DNA methylation and is permissive for binding of transcriptional activators that enable active gene expression ("ON"). Silenced chromatin is characterised by deacetylated histones, enriched with repressive histone modifications and DNA methylation, which lead to binding of repressor complexes and chromatin compaction. Compacted chromatin prevents the binding of transcriptional activators to regulatory elements and contributes to gene repression ("OFF").

Figure 6 DNA methylation profiling using bisulfite-based methods.

Bisulfite conversion of the isolated DNA from biomaterials enables the analysis of DNA methylation in each CpG site, as it allows differentiation of methylated and unmethylated cytosines based on their differential modification upon treatment. Unmethylated cytosines are deaminated to uracils, which get substituted with thymines in the subsequent PCR step. Methylated cytosines are protected from deamination and remain as cytosines. Converted DNA can then be hybridised to a set of probes on an array (e..g on Illumina Epic array) or used for next-generation sequencing library preparation, followed by sequencing (RRBS, WGBS). As a result, differentially methylated probes (DMPs) or differentially methylated regions (DMRs) can be identified and used in biomarker development or further characterised in mechanistic studies.

Figure 7 Targeted DNA methylation and demethylation using CRISPR-based epigenetic editing.

Principle of epigenetic editing for gene activation (top) and gene repression (bottom). A programable DNA targeting domain consisting of the catalytically dead CRISPR/Cas9 (dCas9) is fused to an epigenetic effector domain, like TET demethylase or DNA methyltransferase (DNMT). The fusion domain can be targeted by specific guide RNAs (gRNAs) to a desired genomic region (e.g. gene promoter) via recognition of the target sequence. Once targeted, the epigenetic effector domain can remove (top) or introduce (bottom) DNA methylation modification, leading to subsequent gene silencing (a) or activation (b). gRNA sequences can be designed to target any desired gene/region and a combination of different epigenetic effector domains can be used to target different layers of epigenetic regulation.

Tables

References

Adam, S., Bracker, J., Klingel, V., Osteresch, B., Radde, N. E., Brockmeyer, J., et al. (2022). Flanking sequences influence the activity of TET1 and TET2 methylcytosine dioxygenases and affect genomic 5hmC patterns. *Commun Biol, 5*, 92.

Adams, T. S., Marlier, A., & Kaminski, N. (2023). Lung Cell Atlases in Health and Disease. *Annu Rev Physiol, 85*, 47-69.

Ahn, J., Heo, S., Lee, J., & Bang, D. (2021). Introduction to Single-Cell DNA Methylation Profiling Methods. *Biomolecules, 11*.

Allen, R. J., Porte, J., Braybrooke, R., Flores, C., Fingerlin, T. E., Oldham, J. M., et al. (2017). Genetic variants associated with susceptibility to idiopathic pulmonary fibrosis in people of European ancestry: a genome-wide association study. *Lancet Respir Med, 5*, 869-880.

Allis, C. D., & Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat Rev Genet, 17*, 487-500.

Ambatipudi, S., Cuenin, C., Hernandez-Vargas, H., Ghantous, A., Le Calvez-Kelm, F., Kaaks, R., et al. (2016). Tobacco smoking-associated genome-wide DNA methylation changes in the EPIC study. *Epigenomics, 8*, 599-618.

Antonini, J. M., Kodali, V., Meighan, T. G., Roach, K. A., Roberts, J. R., Salmen, R., et al. (2019). Effect of Age, High-Fat Diet, and Rat Strain on Serum Biomarkers and Telomere Length and Global DNA Methylation in Peripheral Blood Mononuclear Cells. *Sci Rep, 9*, 1996.

Anway, M. D., Cupp, A. S., Uzumcu, M., & Skinner, M. K. (2005). Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science, 308*, 1466-1469.

Ashley, E. A. (2016). Towards precision medicine. *Nat Rev Genet, 17*, 507-522.

Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A., & Balasubramanian, S. (2014). 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat Chem, 6*, 1049-1055.

Barnes, P. J., Burney, P. G., Silverman, E. K., Celli, B. R., Vestbo, J., Wedzicha, J. A., et al. (2015). Chronic obstructive pulmonary disease. *Nat Rev Dis Primers, 1*, 15076.

Bashtrykov, P., Jankevicius, G., Jurkowska, R. Z., Ragozin, S., & Jeltsch, A. (2014). The UHRF1 protein stimulates the activity and specificity of the maintenance DNA methyltransferase DNMT1 by an allosteric mechanism. *J Biol Chem, 289*, 4106-4115.

Baylin, S. B., Hoppener, J. W., de Bustros, A., Steenbergh, P. H., Lips, C. J., & Nelkin, B. D. (1986). DNA methylation patterns of the calcitonin gene in human lung cancers and lymphomas. *Cancer Research, 46*, 2917-2922.

Baysoy, A., Bai, Z., Satija, R., & Fan, R. (2023). The technological landscape and applications of single-cell multi-omics. *Nat Rev Mol Cell Biol, 24*, 695-713.

Beane, J., Sebastiani, P., Liu, G., Brody, J. S., Lenburg, M. E., & Spira, A. (2007). Reversible and permanent effects of tobacco smoke exposure on airway epithelial gene expression. *Genome Biol, 8*, R201.

Belinsky, S. A., Palmisano, W. A., Gilliland, F. D., Crooks, L. A., Divine, K. K., Winters, S. A., et al. (2002). Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. *Cancer Research, 62*, 2370-2377.

Bergman, Y., & Cedar, H. (2013). DNA methylation dynamics in health and disease. *Nat Struct Mol Biol, 20*, 274-281.

Bergstedt, J., Azzou, S. A. K., Tsuo, K., Jaquaniello, A., Urrutia, A., Rotival, M., et al. (2022). The immune factors driving DNA methylation variation in human blood. *Nat Commun, 13*, 5895.

Bermingham, M. L., Walker, R. M., Marioni, R. E., Morris, S. W., Rawlik, K., Zeng, Y., et al. (2019). Identification of novel differentially methylated sites with potential as clinical predictors of impaired respiratory function and COPD. *EBioMedicine, 43*, 576-586.

Bernstein, B. E., Stamatoyannopoulos, J. A., Costello, J. F., Ren, B., Milosavljevic, A., Meissner, A., et al. (2010). The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol, 28*, 1045-1048.

Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J. M., et al. (2011). High density DNA methylation array with single CpG site resolution. *Genomics, 98*, 288-295.

Bibikova, M., Le, J., Barnes, B., Saedinia-Melnyk, S., Zhou, L., Shen, R., et al. (2009). Genome-wide DNA methylation profiling using Infinium(R) assay. *Epigenomics, 1*, 177-200.

Birney, E., Smith, G. D., & Greally, J. M. (2016). Epigenome-wide Association Studies and the Interpretation of Disease -Omics. *PLoS Genet, 12*, e1006105.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., et al. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science, 326*, 1509-1512.

Bojesen, S. E., Timpson, N., Relton, C., Davey Smith, G., & Nordestgaard, B. G. (2017). AHRR(cg05575921) hypomethylation marks smoking behaviour, morbidity and mortality. *Thorax, 72*, 646-653.

Bollepalli, S., Korhonen, T., Kaprio, J., Anders, S., & Ollikainen, M. (2019). EpiSmokEr: a robust classifier to determine smoking status from DNA methylation data. *Epigenomics, 11*, 1469-1486.

Bondarev, A. D., Attwood, M. M., Jonsson, J., Chubarev, V. N., Tarasov, V. V., & Schioth, H. B. (2021). Recent developments of HDAC inhibitors: Emerging indications and novel molecules. *Br J Clin Pharmacol, 87*, 4577-4597.

Bonnelykke, K., & Ober, C. (2016). Leveraging gene-environment interactions and endotypes for asthma gene discovery. *J Allergy Clin Immunol, 137*, 667-679.

Booth, S., Hsieh, A., Mostaco-Guidolin, L., Koo, H. K., Wu, K., Aminazadeh, F., et al. (2023). A Single-Cell Atlas of Small Airway Disease in Chronic Obstructive Pulmonary Disease: A Cross-Sectional Study. *Am J Respir Crit Care Med, 208*, 472-486.

Bosse, Y., Postma, D. S., Sin, D. D., Lamontagne, M., Couture, C., Gaudreault, N., et al. (2012). Molecular signature of smoking in human lung tissues. *Cancer Research, 72*, 3753-3763.

Bostick, M., Kim, J. K., Esteve, P. O., Clark, A., Pradhan, S., & Jacobsen, S. E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science, 317*, 1760-1764.

Buro-Auriemma, L. J., Salit, J., Hackett, N. R., Walters, M. S., Strulovici-Barel, Y., Staudt, M. R., et al. (2013). Cigarette smoking induces small airway epithelial epigenetic changes with corresponding modulation of gene expression. *Human Molecular Genetics, 22*, 4726-4738.

Busch, R., Qiu, W., Lasky-Su, J., Morrow, J., Criner, G., & DeMeo, D. (2016). Differential DNA methylation marks and gene comethylation of COPD in African-Americans with COPD exacerbations. *Respir Res, 17*, 143.

Cardenas, A., Sordillo, J. E., Rifas-Shiman, S. L., Chung, W., Liang, L., Coull, B. A., et al. (2019). The nasal methylome as a biomarker of asthma and airway inflammation in children. *Nat Commun, 10*, 3095.

Carmona, J. J., Barfield, R. T., Panni, T., Nwanaji-Enwerem, J. C., Just, A. C., Hutchinson, J. N., et al. (2018). Metastable DNA methylation sites associated with longitudinal lung function decline and aging in humans: an epigenome-wide study in the NAS and KORA cohorts. *Epigenetics, 13*, 1039-1055.

Casas-Recasens, S., Noell, G., Mendoza, N., Lopez-Giraldo, A., Garcia, T., Guirao, A., et al. (2021). Lung DNA Methylation in Chronic Obstructive Pulmonary Disease: Relationship with Smoking Status and Airflow Limitation Severity. *American Journal of Respiratory and Critical Care Medicine, 203*, 129-134.

Chatziioannou, A., Georgiadis, P., Hebels, D. G., Liampa, I., Valavanis, I., Bergdahl, I. A., et al. (2017). Blood-based omic profiling supports female susceptibility to tobacco smoke-induced cardiovascular diseases. *Sci Rep, 7*, 42870.

Chavez, A., Scheiman, J., Vora, S., Pruitt, B. W., Tuttle, M., E, P. R. I., et al. (2015). Highly efficient Cas9-mediated transcriptional programming. *Nat Methods, 12*, 326-328.

Chen, A., Zeilinger, S., Kühnel, B., Klopp, N., Baurecht, H., Kleinschmidt, A., et al. (2013). Tobacco Smoking Leads to Extensive Genome-Wide Changes in DNA Methylation. *PLoS One, 8*, e63812.

Cho, M. H., McDonald, M. L., Zhou, X., Mattheisen, M., Castaldi, P. J., Hersh, C. P., et al. (2014). Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. *Lancet Respir Med, 2*, 214-225.

Clifford, R. L., Fishbane, N., Patel, J., MacIsaac, J. L., McEwen, L. M., Fisher, A. J., et al. (2018). Altered DNA methylation is associated with aberrant gene expression in parenchymal but not airway fibroblasts isolated from individuals with COPD. *Clin Epigenetics, 10*, 32.

ClinicalTrials.gov. (2022). A phase 1/2 study to evaluate OTX-2022 in patients with hepatocellular carcinoma and other solid tumor types known for association with the MYC oncogene (MYCHELANGELO I). In (Vol. 2024).

Consortium, E. P. (2004). The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science, 306*, 636-640.

Consortium, E. P. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature, 489*, 57-74.

Cosgrove, M. S., Boeke, J. D., & Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. *Nat Struct Mol Biol, 11*, 1037-1043.

Costa-Pinheiro, P., Montezuma, D., Henrique, R., & Jeronimo, C. (2015). Diagnostic and prognostic epigenetic biomarkers in cancer. *Epigenomics, 7*, 1003-1015.

Dao, A., & Bernstein, D. I. (2018). Occupational exposure and asthma. *Ann Allergy Asthma Immunol, 120*, 468-475.

Dogan, M. V., Shields, B., Cutrona, C., Gao, L., Gibbons, F. X., Simons, R., et al. (2014). The effect of smoking on DNA methylation of peripheral blood mononuclear cells from African American women. *BMC Genomics, 15*, 151.

Edris, A., den Dekker, H. T., Melen, E., & Lahousse, L. (2019). Epigenome-wide association studies in asthma: A systematic review. *Clin Exp Allergy, 49*, 953-968.

Ezzati, M., & Lopez, A. D. (2003). Estimates of global mortality attributable to smoking in 2000. *Lancet, 362*, 847-852.

Falkenberg, K. J., & Johnstone, R. W. (2014). Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat Rev Drug Discov, 13*, 673-691.

Fang, F., Andersen, A. M., Philibert, R., & Hancock, D. B. (2023). Epigenetic biomarkers for smoking cessation. *Addict Neurosci, 6*.

Feehley, T., O'Donnell, C. W., Mendlein, J., Karande, M., & McCauley, T. (2023). Drugging the epigenome in the age of precision medicine. *Clin Epigenetics, 15*, 6.

Feinberg, A. P., & Vogelstein, B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature, 301*, 89-92.

Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W. B., Fedorov, O., et al. (2010). Selective inhibition of BET bromodomains. *Nature, 468*, 1067-1073.

Forno, E., Wang, T., Qi, C., Yan, Q., Xu, C. J., Boutaoui, N., et al. (2019). DNA methylation in nasal epithelium, atopy, and atopic asthma in children: a genome-wide study. *Lancet Respir Med, 7*, 336-346.

Franklin, K. A., Shields, C. E., & Haynes, K. A. (2022). Beyond the marks: reader-effectors as drivers of epigenetics and chromatin engineering. *Trends Biochem Sci, 47*, 417-432.

Gama-Sosa, M. A., Slagel, V. A., Trewyn, R. W., Oxenhandler, R., Kuo, K. C., Gehrke, C. W., et al. (1983). The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Res, 11*, 6883-6894.

Gandhi, S. A., Min, B., Fazio, J. C., Johannson, K. A., Steinmaus, C., Reynolds, C. J., et al. (2024). The Impact of Occupational Exposures on the Risk of Idiopathic Pulmonary Fibrosis: A Systematic Review and Meta-Analysis. *Ann Am Thorac Soc, 21*, 486-498.

Ganesan, A., Arimondo, P. B., Rots, M. G., Jeronimo, C., & Berdasco, M. (2019). The timeline of epigenetic drug discovery: from reality to dreams. *Clin Epigenetics, 11*, 174.

Gao, X., Jia, M., Zhang, Y., Breitling, L. P., & Brenner, H. (2015). DNA methylation changes of whole blood cells in response to active smoking exposure in adults: a systematic review of DNA methylation studies. *Clin Epigenetics, 7*, 113.

GOLD. (2023). The Global Strategy for Diagnosis, Management and Prevention of Chronic Obstructive Pulmonary Disease (updated 2023). In.

Greenberg, M. V. C., & Bourc'his, D. (2019). The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol, 20*, 590-607.

Greger, V., Passarge, E., Hopping, W., Messmer, E., & Horsthemke, B. (1989). Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Hum Genet, 83*, 155-158.

Guida, F., Sandanger, T. M., Castagne, R., Campanella, G., Polidoro, S., Palli, D., et al. (2015). Dynamics of smoking-induced genome-wide methylation changes with time since smoking cessation. *Human Molecular Genetics, 24*, 2349-2359.

Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., et al. (2015). Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature, 517*, 640-644.

Hancock, D. B., Eijgelsheim, M., Wilk, J. B., Gharib, S. A., Loehr, L. R., Marciante, K. D., et al. (2010). Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. *Nat Genet, 42*, 45-52.

Handa, V., & Jeltsch, A. (2005). Profound flanking sequence preference of Dnmt3a and Dnmt3b mammalian DNA methyltransferases shape the human epigenome. *J Mol Biol, 348*, 1103-1112.

Hatton, I. A., Galbraith, E. D., Merleau, N. S. C., Miettinen, T. P., Smith, B. M., & Shander, J. A. (2023). The human cell count and size distribution. *Proc Natl Acad Sci U S A, 120*, e2303077120.

Heinbockel, L., Marwitz, S., Schromm, A. B., Watz, H., Kugler, C., Ammerpohl, O., et al. (2018). Identification of novel target genes in human lung tissue involved in chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis, 13*, 2255-2259.

Helling, B. A., Gerber, A. N., Kadiyala, V., Sasse, S. K., Pedersen, B. S., Sparks, L., et al. (2017). Regulation of MUC5B Expression in Idiopathic Pulmonary Fibrosis. *Am J Respir Cell Mol Biol, 57*, 91-99.

Hoang, T. T., Lee, Y., McCartney, D. L., Kersten, E. T. G., Page, C. M., Hulls, P. M., et al. (2024). Comprehensive evaluation of smoking exposures and their interactions on DNA methylation. *EBioMedicine, 100*, 104956.

Hobbs, B. D., de Jong, K., Lamontagne, M., Bosse, Y., Shrine, N., Artigas, M. S., et al. (2017). Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nat Genet, 49*, 426-432.

Hobbs, B. D., Morrow, J. D., Wang, X.-W., Liu, Y.-Y., DeMeo, D. L., Hersh, C. P., et al. (2023). Identifying chronic obstructive pulmonary disease from integrative omics and clustering in lung tissue. *BMC Pulmonary Medicine, 23*, 115.

Horii, T., Morita, S., Hino, S., Kimura, M., Hino, Y., Kogo, H., et al. (2020). Successful generation of epigenetic disease model mice by targeted demethylation of the epigenome. *Genome Biol, 21*, 77.

Horvath, S., & Raj, K. (2018). DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet, 19*, 371-384.

Houseman, E. A., Kelsey, K. T., Wiencke, J. K., & Marsit, C. J. (2015). Cell-composition effects in the analysis of DNA methylation array data: a mathematical perspective. *BMC Bioinformatics, 16*, 95.

Hsu, P. D., Lander, E. S., & Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell, 157*, 1262-1278.

Hu, C., Liu, X., Zeng, Y., Liu, J., & Wu, F. (2021). DNA methyltransferase inhibitors combination therapy for the treatment of solid tumor: mechanism and clinical application. *Clin Epigenetics, 13*, 166.

Huang, S. K., Fisher, A. S., Scruggs, A. M., White, E. S., Hogaboam, C. M., Richardson, B. C., et al. (2010). Hypermethylation of PTGER2 confers prostaglandin E2 resistance in fibrotic fibroblasts from humans and mice. *Am J Pathol, 177*, 2245-2255.

Huang, S. K., Scruggs, A. M., McEachin, R. C., White, E. S., & Peters-Golden, M. (2014). Lung fibroblasts from patients with idiopathic pulmonary fibrosis exhibit genome-wide differences in DNA methylation compared to fibroblasts from nonfibrotic lung. *PLoS One, 9*, e107055.

Huen, K., Yousefi, P., Bradman, A., Yan, L., Harley, K. G., Kogut, K., et al. (2014). Effects of age, sex, and persistent organic pollutants on DNA methylation in children. *Environ Mol Mutagen, 55*, 209-222.

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., et al. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science, 333*, 1300-1303.

Jaffe, A. E., & Irizarry, R. A. (2014). Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol, 15*, R31.

Jeltsch, A. (2013). Oxygen, epigenetic signaling, and the evolution of early life. *Trends Biochem Sci, 38*, 172-176.

Jeltsch, A., & Jurkowska, R. Z. (2013). Multimerization of the dnmt3a DNA methyltransferase and its functional implications. *Prog Mol Biol Transl Sci, 117*, 445-464.

Jeltsch, A., & Jurkowska, R. Z. (2014). New concepts in DNA methylation. *Trends Biochem Sci, 39*, 310-318.

Jeltsch, A., & Jurkowska, R. Z. (2016). Allosteric control of mammalian DNA methyltransferases - a new regulatory paradigm. *Nucleic Acids Res, 44*, 8556-8575.

Jeltsch, A., Jurkowska, R. Z., Jurkowski, T. P., Liebert, K., Rathert, P., & Schlickenrieder, M. (2007). Application of DNA methyltransferases in targeted DNA methylation. *Appl Microbiol Biotechnol, 75*, 1233-1240.

Jeong, Y., de Andrade, E. S. L. B., Thalmeier, D., Toth, R., Ganslmeier, M., Breuer, K., et al. (2022). Systematic evaluation of cell-type deconvolution pipelines for sequencing-based bulk DNA methylomes. *Brief Bioinform, 23*.

Jeziorska, D. M., Murray, R. J. S., De Gobbi, M., Gaentzsch, R., Garrick, D., Ayyub, H., et al. (2017). DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. *Proc Natl Acad Sci U S A, 114*, E7526-E7535.

Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A., & Cheng, X. (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature, 449*, 248-251.

Joehanes, R., Just, A. C., Marioni, R. E., Pilling, L. C., Reynolds, L. M., Mandaviya, P. R., et al. (2016). Epigenetic Signatures of Cigarette Smoking. *Circ Cardiovasc Genet, 9*, 436-447.

Jones, P. A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet, 13*, 484-492.

Jones, P. A., & Taylor, S. M. (1980). Cellular differentiation, cytidine analogs and DNA methylation. *Cell, 20*, 85-93.

Joubert, B. R., Haberg, S. E., Nilsen, R. M., Wang, X., Vollset, S. E., Murphy, S. K., et al. (2012). 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect, 120*, 1425-1431.

Jurkowska, R. Z., & Jeltsch, A. (2022a). Enzymology of Mammalian DNA Methyltransferases. *Adv Exp Med Biol, 1389*, 69-110.

Jurkowska, R. Z., & Jeltsch, A. (2022b). Mechanisms and Biological Roles of DNA Methyltransferases and DNA Methylation: From Past Achievements to Future Challenges. *Adv Exp Med Biol, 1389*, 1-19.

Jurkowska, R. Z., & Jurkowski, T. P. (2019). Establishment, Erasure and Synthetic Reprogramming of DNA Methylation in Mammalian Cells. In S. Jurga & J. Barciszewski (Eds.), *The DNA, RNA, and Histone Methylomes* (pp. 1-26). Cham: Springer International Publishing.

Jurkowska, R. Z., Jurkowski, T. P., & Jeltsch, A. (2011). Structure and function of mammalian DNA methyltransferases. *Chembiochem, 12*, 206-222.

Jurkowska, R. Z., Rajavelu, A., Anspach, N., Urbanke, C., Jankevicius, G., Ragozin, S., et al. (2011). Oligomerization and binding of the Dnmt3a DNA methyltransferase to parallel DNA molecules: heterochromatic localization and role of Dnmt3L. *J Biol Chem, 286*, 24200-24207.

Jurkowska, R. Z., Siddique, A. N., Jurkowski, T. P., & Jeltsch, A. (2011). Approaches to enzyme and substrate design of the murine Dnmt3a DNA methyltransferase. *Chembiochem, 12*, 1589-1594.

Jurkowski, T. P., Ravichandran, M., & Stepper, P. (2015). Synthetic epigenetics-towards intelligent control of epigenetic states and cell identity. *Clin Epigenetics, 7*, 18.

Kachroo, P., Morrow, J. D., Kho, A. T., Vyhlidal, C. A., Silverman, E. K., Weiss, S. T., et al. (2020). Co-methylation analysis in lung tissue identifies pathways for fetal origins of COPD. *European Respiratory Journal, 56*.

Kaur, D., Lee, S. M., Goldberg, D., Spix, N. J., Hinoue, T., Li, H.-T., et al. (2023). Comprehensive evaluation of the Infinium human MethylationEPIC v2 BeadChip. *Epigenetics Communications, 3*, 6.

Kaur, G., Begum, R., Thota, S., & Batra, S. (2019). A systematic review of smoking-related epigenetic alterations. *Archives of Toxicology, 93*, 2715-2740.

Keshet, I., Lieman-Hurwitz, J., & Cedar, H. (1986). DNA methylation affects the formation of active chromatin. *Cell, 44*, 535-543.

Laird, P. W. (2010). Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet, 11*, 191-203.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature, 409*, 860-921.

Lau, C. H., & Suh, Y. (2018). In vivo epigenome editing and transcriptional modulation using CRISPR technology. *Transgenic Res, 27*, 489-509.

Lederer, D. J., & Martinez, F. J. (2018). Idiopathic Pulmonary Fibrosis. *N Engl J Med, 378*, 1811-1823.

Lee, J. U., Son, J. H., Shim, E. Y., Cheong, H. S., Shin, S. W., Shin, H. D., et al. (2019). Global DNA Methylation Pattern of Fibroblasts in Idiopathic Pulmonary Fibrosis. *DNA Cell Biol, 38*, 905-914.

Legaki, E., Arsenis, C., Taka, S., & Papadopoulos, N. G. (2022). DNA methylation biomarkers in asthma and rhinitis: Are we there yet? *Clin Transl Allergy, 12*, e12131.

Li, E., Bestor, T. H., & Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell, 69*, 915-926.

Liao, H. K., Hatanaka, F., Araoka, T., Reddy, P., Wu, M. Z., Sui, Y., et al. (2017). In Vivo Target Gene Activation via CRISPR/Cas9-Mediated Trans-epigenetic Modulation. *Cell, 171*, 1495-1507 e1415.

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature, 462*, 315-322.

Liu, F., Killian, J. K., Yang, M., Walker, R. L., Hong, J. A., Zhang, M., et al. (2010). Epigenomic alterations and gene expression profiles in respiratory epithelia exposed to cigarette smoke condensate. *Oncogene, 29*, 3650-3664.

Liu, R., Zhao, E., Yu, H., Yuan, C., Abbas, M. N., & Cui, H. (2023). Methylation across the central dogma in health and diseases: new therapeutic strategies. *Signal Transduct Target Ther, 8*, 310.

Liu, X. S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., et al. (2016). Editing DNA Methylation in the Mammalian Genome. *Cell, 167*, 233-247 e217.

Liu, Y., Hu, Z., Cheng, J., Siejka-Zielinska, P., Chen, J., Inoue, M., et al. (2021). Subtraction-free and bisulfite-free specific sequencing of 5-methylcytosine and its oxidized derivatives at base resolution. *Nat Commun, 12*, 618.

Liu, Y., Siejka-Zielinska, P., Velikova, G., Bi, Y., Yuan, F., Tomkova, M., et al. (2019). Bisulfite-free direct detection of 5-methylcytosine and 5-hydroxymethylcytosine at base resolution. *Nat Biotechnol, 37*, 424-429.
Lizio, M., Harshbarger, J., Shimoji, H., Severin, J., Kasukawa, T., Sahin, S., et al. (2015). Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biol, 16*, 22.
Llamazares-Prada, M., Espinet, E., Mijosek, V., Schwartz, U., Lutsik, P., Tamas, R., et al. (2021). Versatile workflow for cell type-resolved transcriptional and epigenetic profiles from cryopreserved human lung. *JCI Insight, 6*.
Loda, A., & Heard, E. (2019). Xist RNA in action: Past, present, and future. *PLoS Genet, 15*, e1008333.
Loyfer, N., Magenheim, J., Peretz, A., Cann, G., Bredno, J., Klochendler, A., et al. (2023). A DNA methylation atlas of normal human cell types. *Nature, 613*, 355-364.
Ma, J., & Ge, Z. (2021). Comparison Between Decitabine and Azacitidine for Patients With Acute Myeloid Leukemia and Higher-Risk Myelodysplastic Syndrome: A Systematic Review and Network Meta-Analysis. *Front Pharmacol, 12*, 701690.
Maas, S. C. E., Vidaki, A., Wilson, R., Teumer, A., Liu, F., van Meurs, J. B. J., et al. (2019). Validated inference of smoking habits from blood with a finite DNA methylation marker set. *Eur J Epidemiol, 34*, 1055-1074.
Martens, J. H., & Stunnenberg, H. G. (2013). BLUEPRINT: mapping human blood cell epigenomes. *Haematologica, 98*, 1487-1489.
Martin, E. M., & Fry, R. C. (2018). Environmental Influences on the Epigenome: Exposure- Associated DNA Methylation in Human Populations. *Annu Rev Public Health, 39*, 309-333.
Martinez, F. J., Chisholm, A., Collard, H. R., Flaherty, K. R., Myers, J., Raghu, G., et al. (2017). The diagnosis of idiopathic pulmonary fibrosis: current and future approaches. *Lancet Respir Med, 5*, 61-71.
Matharu, N., Rattanasopha, S., Tamura, S., Maliskova, L., Wang, Y., Bernard, A., et al. (2019). CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. *Science, 363*.
Matthews, G. M., Newbold, A., & Johnstone, R. W. (2012). Intrinsic and extrinsic apoptotic pathway signaling as determinants of histone deacetylase inhibitor antitumor activity. *Adv Cancer Res, 116*, 165-197.
Mattick, J. S., Amaral, P. P., Carninci, P., Carpenter, S., Chang, H. Y., Chen, L.-L., et al. (2023). Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nature Reviews Molecular Cell Biology, 24*, 430-447.
McCartney, D. L., Hillary, R. F., Stevenson, A. J., Ritchie, S. J., Walker, R. M., Zhang, Q., et al. (2018). Epigenetic prediction of complex traits and death. *Genome Biol, 19*, 136.
McCartney, D. L., Stevenson, A. J., Hillary, R. F., Walker, R. M., Bermingham, M. L., Morris, S. W., et al. (2018). Epigenetic signatures of starting and stopping smoking. *EBioMedicine, 37*, 214-220.
McErlean, P., Bell, C. G., Hewitt, R. J., Busharat, Z., Ogger, P. P., Ghai, P., et al. (2021). DNA Methylome Alterations Are Associated with Airway Macrophage Differentiation and Phenotype during Lung Fibrosis. *Am J Respir Crit Care Med, 204*, 954-966.
Meissner, A., Gnirke, A., Bell, G. W., Ramsahoye, B., Lander, E. S., & Jaenisch, R. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res, 33*, 5868-5877.
Melen, E., Koppelman, G. H., Vicedo-Cabrera, A. M., Andersen, Z. J., & Bunyavanich, S. (2022). Allergies to food and airborne allergens in children and adolescents: role of epigenetics in a changing environment. *Lancet Child Adolesc Health, 6*, 810-819.
Michalak, E. M., Burr, M. L., Bannister, A. J., & Dawson, M. A. (2019). The roles of DNA, RNA and histone methylation in ageing and cancer. *Nat Rev Mol Cell Biol, 20*, 573-589.
Millán-Zambrano, G., Burton, A., Bannister, A. J., & Schneider, R. (2022). Histone post-translational modifications — cause and consequence of genome function. *Nature Reviews Genetics, 23*, 563-580.

Montoro, D. T., Haber, A. L., Biton, M., Vinarsky, V., Lin, B., Birket, S. E., et al. (2018). A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature, 560*, 319-324.
Moran, S., Vizoso, M., Martinez-Cardus, A., Gomez, A., Matias-Guiu, X., Chiavenna, S. M., et al. (2014). Validation of DNA methylation profiling in formalin-fixed paraffin-embedded samples using the Infinium HumanMethylation450 Microarray. *Epigenetics, 9*, 829-833.
Morel, D., Jeffery, D., Aspeslagh, S., Almouzni, G., & Postel-Vinay, S. (2020). Combining epigenetic drugs with other therapies for solid tumours - past lessons and future promise. *Nat Rev Clin Oncol, 17*, 91-107.
Morrow, J. D., Cho, M. H., Hersh, C. P., Pinto-Plata, V., Celli, B., Marchetti, N., et al. (2016). DNA methylation profiling in human lung tissue identifies genes associated with COPD. *Epigenetics, 11*, 730-739.
Murgia, N., & Gambelunghe, A. (2022). Occupational COPD-The most under-recognized occupational lung disease? *Respirology, 27*, 399-410.
Nakamura, M., Gao, Y., Dominguez, A. A., & Qi, L. S. (2021). CRISPR technologies for precise epigenome editing. *Nat Cell Biol, 23*, 11-22.
Negretti, N. M., Plosa, E. J., Benjamin, J. T., Schuler, B. A., Habermann, A. C., Jetter, C. S., et al. (2021). A single-cell atlas of mouse lung development. *Development, 148*.
Nehme, Z., Pasquereau, S., & Herbein, G. (2019). Control of viral infections by epigenetic-targeted therapy. *Clin Epigenetics, 11*, 55.
Nemeth, K., Bayraktar, R., Ferracin, M., & Calin, G. A. (2024). Non-coding RNAs in disease: from mechanisms to therapeutics. *Nature Reviews Genetics, 25*, 211-232.
Nicodeme, E., Jeffrey, K. L., Schaefer, U., Beinke, S., Dewell, S., Chung, C. W., et al. (2010). Suppression of inflammation by a synthetic histone mimic. *Nature, 468*, 1119-1123.
Nicodemus-Johnson, J., Myers, R. A., Sakabe, N. J., Sobreira, D. R., Hogarth, D. K., Naureckas, E. T., et al. (2016). DNA methylation in lung cells is associated with asthma endotypes and genetic risk. *JCI Insight, 1*, e90151.
Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A. V., Mikheenko, A., et al. (2022). The complete sequence of a human genome. *Science, 376*, 44-53.
Nwanaji-Enwerem, J. C., & Colicino, E. (2020). DNA Methylation-Based Biomarkers of Environmental Exposures for Human Population Studies. *Curr Environ Health Rep, 7*, 121-128.
Okano, M., Bell, D. W., Haber, D. A., & Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell, 99*, 247-257.
Olstad, E. W., Nordeng, H. M. E., Sandve, G. K., Lyle, R., & Gervin, K. (2022). Low reliability of DNA methylation across Illumina Infinium platforms in cord blood: implications for replication studies and meta-analyses of prenatal exposures. *Clin Epigenetics, 14*, 80.
Olvera, N., Sanchez-Valle, J., Nunez-Carpintero, I., Rojas-Quintero, J., Noell, G., Casas-Recasens, S., et al. (2024). Lung Tissue Multi-Layer Network Analysis Uncovers the Molecular Heterogeneity of COPD. *Am J Respir Crit Care Med*.
Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., & Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO Rep, 10*, 1235-1241.
Parker, M. M., Chase, R. P., Lamb, A., Reyes, A., Saferali, A., Yun, J. H., et al. (2017). RNA sequencing identifies novel non-coding RNA and exon-specific effects associated with cigarette smoking. *BMC Med Genomics, 10*, 58.
Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods, 10*, 973-976.
Petryk, N., Bultmann, S., Bartke, T., & Defossez, P. A. (2021). Staying true to yourself: mechanisms of DNA methylation maintenance in mammals. *Nucleic Acids Res, 49*, 3020-3032.
Philibert, R., Monick, M. M., Beach, S. R. H., Plume, J., Sears, R., Gerrard, M., et al. (2012). Smoking is Associated with Coordinated Changes in Aryl Hydrocarbon Receptor Repressor Methylation

in Genome Wide Analyses of Lymphoblast and Pulmonary Macrophage DNA. *Biological Psychiatry, 71*, 163s-163s.
Philibert, R. A., Beach, S. R., Lei, M. K., & Brody, G. H. (2013). Changes in DNA methylation at the aryl hydrocarbon receptor repressor may be a new biomarker for smoking. *Clin Epigenetics, 5*, 19.
Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy, P., et al. (2016). Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol, 17*, 208.
Plasschaert, L. W., Zilionis, R., Choo-Wing, R., Savova, V., Knehr, J., Roma, G., et al. (2018). A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature, 560*, 377-381.
Pohl, S. T., Prada, M. L., Espinet, E., & Jurkowska, R. (2023). Practical Considerations for Complex Tissue Dissociation for Single-Cell Transcriptomics. *Methods Mol Biol, 2584*, 371-387.
Polverino, F., & Kalhan, R. (2023). Leveraging Omics to Predict Chronic Obstructive Pulmonary Disease Exacerbations: The "Immunome". *Am J Respir Crit Care Med, 208*, 220-222.
Porsbjerg, C., Melén, E., Lehtimäki, L., & Shaw, D. (2023). Asthma. *The Lancet, 401*, 858-873.
Prada, M. L., Schwartz, U., Pease, D. F., Pohl, S. T., Ackesson, D., Li, R., et al. (2023). Epigenetic deregulation of IFN and WNT pathways in AT2 cells impairs alveolar regeneration (in COPD). *bioRxiv*, 2023.2010.2022.563483.
Prevention, C. f. D. C. a. (2014). *National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health. The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General.* Atlanta (GA): .
Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., et al. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell, 152*, 1173-1183.
Qiu, W., Baccarelli, A., Carey, V. J., Boutaoui, N., Bacherman, H., Klanderman, B., et al. (2012). Variable DNA methylation is associated with chronic obstructive pulmonary disease and lung function. *Am J Respir Crit Care Med, 185*, 373-381.
Rabinovich, E. I., Kapetanaki, M. G., Steinfeld, I., Gibson, K. F., Pandit, K. V., Yu, G., et al. (2012). Global methylation patterns in idiopathic pulmonary fibrosis. *PLoS One, 7*, e33770.
Raghu, G., Remy-Jardin, M., Myers, J. L., Richeldi, L., Ryerson, C. J., Lederer, D. J., et al. (2018). Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. *Am J Respir Crit Care Med, 198*, e44-e68.
Ragland, M. F., Benway, C. J., Lutz, S. M., Bowler, R. P., Hecker, J., Hokanson, J. E., et al. (2019). Genetic Advances in Chronic Obstructive Pulmonary Disease. Insights from COPDGene. *Am J Respir Crit Care Med, 200*, 677-690.
Rajavelu, A., Lungu, C., Emperle, M., Dukatz, M., Brohm, A., Broche, J., et al. (2018). Chromatin-dependent allosteric regulation of DNMT3A activity by MeCP2. *Nucleic Acids Res, 46*, 9044-9056.
Ravichandran, M., Jurkowska, R. Z., & Jurkowski, T. P. (2018). Target specificity of mammalian DNA methylation and demethylation machinery. *Org Biomol Chem, 16*, 1419-1435.
Ravichandran, M., Rafalski, D., Davies, C. I., Ortega-Recalde, O., Nan, X., Glanfield, C. R., et al. (2022). Pronounced sequence specificity of the TET enzyme catalytic domain guides its cellular function. *Sci Adv, 8*, eabm2427.
Reese, S. E., Xu, C. J., den Dekker, H. T., Lee, M. K., Sikdar, S., Ruiz-Arenas, C., et al. (2019). Epigenome-wide meta-analysis of DNA methylation and childhood asthma. *J Allergy Clin Immunol, 143*, 2062-2074.
Richmond, R. C., Simpkin, A. J., Woodward, G., Gaunt, T. R., Lyttleton, O., McArdle, W. L., et al. (2015). Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Human Molecular Genetics, 24*, 2201-2217.

Riggs, A., Porter, TN. (1996). Overview of epgenetic mechanisms. In C. S. H. Cold Spring Harbor Laboratory Press, NY (Ed.), *Epigenetic mechanisms of gene regulation* (pp. 29-45).

Rodriguez-Rivera, I., Wu, T., Ciotti, R., Senapedis, W., Sullivan, K., Gao, J. Z., et al. (2023). A phase 1/2 open-label study to evaluate the safety, tolerability, pharmacokinetics, pharmacodynamics, and preliminary antitumor activity of OTX-2002 as a single agent and in combination with standard of care in patients with hepatocellular carcinoma and other solid tumor types known for association with the MYC oncogene (MYCHELANGELO I). *Journal of Clinical Oncology, 41*

Ruiz Garcia, S., Deprez, M., Lebrigand, K., Cavard, A., Paquet, A., Arguel, M. J., et al. (2019). Novel dynamics of human mucociliary differentiation revealed by single-cell RNA sequencing of nasal epithelial cultures. *Development, 146*.

Sakornsakolpat, P., Prokopenko, D., Lamontagne, M., Reeve, N. F., Guyatt, A. L., Jackson, V. E., et al. (2019). Genetic landscape of chronic obstructive pulmonary disease identifies heterogeneous cell-type and phenotype associations. *Nat Genet, 51*, 494-505.

Sanders, Y. Y., Ambalavanan, N., Halloran, B., Zhang, X., Liu, H., Crossman, D. K., et al. (2012). Altered DNA methylation profile in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med, 186*, 525-535.

Sanders, Y. Y., Liu, H., Scruggs, A. M., Duncan, S. R., Huang, S. K., & Thannickal, V. J. (2017). Epigenetic Regulation of Caveolin-1 Gene Expression in Lung Fibroblasts. *Am J Respir Cell Mol Biol, 56*, 50-61.

Sanders, Y. Y., Pardo, A., Selman, M., Nuovo, G. J., Tollefsbol, T. O., Siegal, G. P., et al. (2008). Thy-1 promoter hypermethylation: a novel epigenetic pathogenic mechanism in pulmonary fibrosis. *Am J Respir Cell Mol Biol, 39*, 610-618.

Saunderson, E. A., Stepper, P., Gomm, J. J., Hoa, L., Morgan, A., Allen, M. D., et al. (2017). Hit-and-run epigenetic editing prevents senescence entry in primary breast cells from healthy donors. *Nat Commun, 8*, 1450.

Schubeler, D. (2015). Function and information content of DNA methylation. *Nature, 517*, 321-326.

Schultz, M. D., He, Y., Whitaker, J. W., Hariharan, M., Mukamel, E. A., Leung, D., et al. (2015). Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature, 523*, 212-216.

Schutsky, E. K., DeNizio, J. E., Hu, P., Liu, M. Y., Nabel, C. S., Fabyanic, E. B., et al. (2018). Nondestructive, base-resolution sequencing of 5-hydroxymethylcytosine using a DNA deaminase. *Nat Biotechnol*.

Schwartz, U., Llamazares Prada, M., Pohl, S. T., Richter, M., Tamas, R., Schuler, M., et al. (2023). High-resolution transcriptomic and epigenetic profiling identifies novel regulators of COPD. *EMBO J, 42*, e111272.

Scruggs, A. M., Koh, H. B., Tripathi, P., Leeper, N. J., White, E. S., & Huang, S. K. (2018). Loss of CDKN2B Promotes Fibrosis via Increased Fibroblast Differentiation Rather Than Proliferation. *Am J Respir Cell Mol Biol, 59*, 200-214.

Seiler, C. L., Song, J. U. M., Kotandeniya, D., Chen, J. J., Kono, T. J. Y., Han, Q. Y., et al. (2020). Inhalation exposure to cigarette smoke and inflammatory agents induces epigenetic changes in the lung. *Scientific Reports, 10*.

Sheikhpour, M., Maleki, M., Ebrahimi Vargoorani, M., & Amiri, V. (2021). A review of epigenetic changes in asthma: methylation and acetylation. *Clin Epigenetics, 13*, 65.

Shenker, N. S., Polidoro, S., van Veldhoven, K., Sacerdote, C., Ricceri, F., Birrell, M. A., et al. (2013a). Epigenome-wide association study in the European Prospective Investigation into Cancer and Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. *Human Molecular Genetics, 22*, 843-851.

Shenker, N. S., Polidoro, S., van Veldhoven, K., Sacerdote, C., Ricceri, F., Birrell, M. A., et al. (2013b). Epigenome-wide association study in the European Prospective Investigation into Cancer and

Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. *Human Molecular Genetics, 22*, 843-851.
Shoeb, M., Mustafa, G. M., Kodali, V. K., Smith, K., Roach, K. A., Boyce, G., et al. (2020). A possible relationship between telomere length and markers of neurodegeneration in rat brain after welding fume inhalation exposure. *Environ Res, 180*, 108900.
Silva, C. P., & Kamens, H. M. (2021). Cigarette smoke-induced alterations in blood: A review of research on DNA methylation and gene expression. *Exp Clin Psychopharmacol, 29*, 116-135.
Singer, B. D. (2019). A Practical Guide to the Measurement and Analysis of DNA Methylation. *Am J Respir Cell Mol Biol, 61*, 417-428.
Smith, Z. D., Hetzel, S., & Meissner, A. (2024). DNA methylation in mammalian development and disease. *Nat Rev Genet*.
Smith, Z. D., & Meissner, A. (2013). DNA methylation: roles in mammalian development. *Nat Rev Genet, 14*, 204-220.
Solazzo, G., Ferrante, G., & La Grutta, S. (2020). DNA Methylation in Nasal Epithelium: Strengths and Limitations of an Emergent Biomarker for Childhood Asthma. *Front Pediatr, 8*, 256.
Song, J., Cano-Rodriquez, D., Winkle, M., Gjaltema, R. A., Goubert, D., Jurkowski, T. P., et al. (2017). Targeted epigenetic editing of SPDEF reduces mucus production in lung epithelial cells. *Am J Physiol Lung Cell Mol Physiol, 312*, L334-L347.
Song, J., & Kuan, P. F. (2022). A systematic assessment of cell type deconvolution algorithms for DNA methylation data. *Brief Bioinform, 23*.
Sood, A., Petersen, H., Blanchette, C. M., Meek, P., Picchi, M. A., Belinsky, S. A., et al. (2010). Wood smoke exposure and gene promoter methylation are associated with increased risk for COPD in smokers. *Am J Respir Crit Care Med, 182*, 1098-1104.
Spainhour, J. C., Lim, H. S., Yi, S. V., & Qiu, P. (2019). Correlation Patterns Between DNA Methylation and Gene Expression in The Cancer Genome Atlas. *Cancer Inform, 18*, 1176935119828776.
Stadler, M. B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Schöler, A., et al. (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature, 480*, 490-495.
Stepper, P., Kungulovski, G., Jurkowska, R. Z., Chandra, T., Krueger, F., Reinhardt, R., et al. (2017). Efficient targeted DNA methylation with chimeric dCas9-Dnmt3a-Dnmt3L methyltransferase. *Nucleic Acids Res, 45*, 1703-1713.
Stewart, S. K., Morris, T. J., Guilhamon, P., Bulstrode, H., Bachman, M., Balasubramanian, S., et al. (2015). oxBS-450K: a method for analysing hydroxymethylation using 450K BeadChips. *Methods, 72*, 9-15.
Stunnenberg, H. G., International Human Epigenome, C., & Hirst, M. (2016). The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. *Cell, 167*, 1145-1149.
Sugden, K., Hannon, E. J., Arseneault, L., Belsky, D. W., Corcoran, D. L., Fisher, H. L., et al. (2020). Patterns of Reliability: Assessing the Reproducibility and Integrity of DNA Methylation Measurement. *Patterns (N Y), 1*.
Sundar, I. K., Yin, Q., Baier, B. S., Yan, L., Mazur, W., Li, D., et al. (2017). DNA methylation profiling in peripheral lung tissues of smokers and patients with COPD. *Clin Epigenetics, 9*, 38.
Svoboda, L. K., Perera, B. P. U., Morgan, R. K., Polemi, K. M., Pan, J., & Dolinoy, D. C. (2022). Toxicoepigenetics and Environmental Health: Challenges and Opportunities. *Chem Res Toxicol, 35*, 1293-1311.
Szyf, M. (2011). The implications of DNA methylation for toxicology: toward toxicomethylomics, the toxicology of DNA methylation. *Toxicol Sci, 120*, 235-255.
Teschendorff, A. E., & Zheng, S. C. (2017). Cell-type deconvolution in epigenome-wide association studies: a review and recommendations. *Epigenomics, 9*, 757-768.
Teschendorff, A. E., Zhu, T., Breeze, C. E., & Beck, S. (2020). EPISCORE: cell type deconvolution of bulk tissue DNA methylomes from single-cell RNA-Seq data. *Genome Biol, 21*, 221.

Thangavel, J., Malik, A. B., Elias, H. K., Rajasingh, S., Simpson, A. D., Sundivakkam, P. K., et al. (2014). Combinatorial therapy with acetylation and methylation modifiers attenuates lung vascular hyperpermeability in endotoxemia-induced mouse inflammatory lung injury. *Am J Pathol, 184*, 2237-2249.

Thienpont, B., Steinbacher, J., Zhao, H., D'Anna, F., Kuchnio, A., Ploumakis, A., et al. (2016). Tumour hypoxia causes DNA hypermethylation by reducing TET activity. *Nature, 537*, 63-68.

Thurmann, L., Klos, M., Mackowiak, S. D., Bieg, M., Bauer, T., Ishaque, N., et al. (2023). Global hypomethylation in childhood asthma identified by genome-wide DNA-methylation sequencing preferentially affects enhancer regions. *Allergy, 78*, 1489-1506.

Titus, A. J., Gallimore, R. M., Salas, L. A., & Christensen, B. C. (2017). Cell-type deconvolution from DNA methylation: a review of recent applications. *Human Molecular Genetics, 26*, R216-R224.

Travaglini, K. J., Nabhan, A. N., Penland, L., Sinha, R., Gillich, A., Sit, R. V., et al. (2020). A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature, 587*, 619-625.

Turner, B. M. (1993). Decoding the nucleosome. *Cell, 75*, 5-8.

Uszczynska-Ratajczak, B., Lagarde, J., Frankish, A., Guigo, R., & Johnson, R. (2018). Towards a complete map of the human long non-coding RNA transcriptome. *Nat Rev Genet, 19*, 535-548.

Vaisvila, R., Ponnaluri, V. K. C., Sun, Z., Langhorst, B. W., Saleh, L., Guan, S., et al. (2021). Enzymatic methyl sequencing detects DNA methylation at single-base resolution from picograms of DNA. *Genome Res, 31*, 1280-1289.

Van de Sande, B., Lee, J. S., Mutasa-Gottgens, E., Naughton, B., Bacon, W., Manning, J., et al. (2023). Applications of single-cell RNA sequencing in drug discovery and development. *Nat Rev Drug Discov, 22*, 496-520.

Varley, K. E., Gertz, J., Bowling, K. M., Parker, S. L., Reddy, T. E., Pauli-Behn, F., et al. (2013). Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res, 23*, 555-567.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., et al. (2001). The sequence of the human genome. *Science, 291*, 1304-1351.

Vogel, C. F. A., & Haarmann-Stemmann, T. (2017). The aryl hydrocarbon receptor repressor - More than a simple feedback inhibitor of AhR signaling: Clues for its role in inflammation and cancer. *Curr Opin Toxicol, 2*, 109-119.

Vucic, E. A., Chari, R., Thu, K. L., Wilson, I. M., Cotton, A. M., Kennett, J. Y., et al. (2014). DNA methylation is globally disrupted and associated with expression changes in chronic obstructive pulmonary disease small airways. *Am J Respir Cell Mol Biol, 50*, 912-922.

Wain, L. V., Shrine, N., Miller, S., Jackson, V. E., Ntalla, I., Soler Artigas, M., et al. (2015). Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir Med, 3*, 769-781.

Walters, G. I. (2020). Occupational exposures and idiopathic pulmonary fibrosis. *Curr Opin Allergy Clin Immunol, 20*, 103-111.

Wan, E. S., Qiu, W., Carey, V. J., Morrow, J., Bacherman, H., Foreman, M. G., et al. (2015). Smoking-Associated Site-Specific Differential Methylation in Buccal Mucosa in the COPDGene Study. *Am J Respir Cell Mol Biol, 53*, 246-254.

Weber, M., Hellmann, I., Stadler, M. B., Ramos, L., Paabo, S., Rebhan, M., et al. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet, 39*, 457-466.

Wei, A., Gao, Q., Chen, F., Zhu, X., Chen, X., Zhang, L., et al. (2022). Inhibition of DNA methylation de-represses peroxisome proliferator-activated receptor-gamma and attenuates pulmonary fibrosis. *Br J Pharmacol, 179*, 1304-1318.

Wiklund, P., Karhunen, V., Richmond, R. C., Parmar, P., Rodriguez, A., De Silva, M., et al. (2019). DNA methylation links prenatal smoking exposure to later life health outcomes in offspring. *Clin Epigenetics, 11*, 97.

Wolfe, S. A., Nekludova, L., & Pabo, C. O. (2000). DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct, 29*, 183-212.

Wu, D.-D., Song, J., Bartel, S., Krauss-Etschmann, S., Rots, M. G., & Hylkema, M. N. (2018). The potential for targeted rewriting of epigenetic marks in COPD as a new therapeutic approach. *Pharmacology & Therapeutics, 182*, 1-14.

Wu, X., & Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet, 18*, 517-534.

Wyss, A. B., Sofer, T., Lee, M. K., Terzikhan, N., Nguyen, J. N., Lahousse, L., et al. (2018). Multiethnic meta-analysis identifies ancestry-specific and cross-ancestry loci for pulmonary function. *Nat Commun, 9*, 2976.

Xu, C. J., Soderhall, C., Bustamante, M., Baiz, N., Gruzieva, O., Gehring, U., et al. (2018). DNA methylation in childhood asthma: an epigenome-wide meta-analysis. *Lancet Respir Med, 6*, 379-388.

Xu, G. L., & Bestor, T. H. (1997). Cytosine methylation targetted to pre-determined sequences. *Nat Genet, 17*, 376-378.

Yang, I. V., Pedersen, B. S., Rabinovich, E., Hennessy, C. E., Davidson, E. J., Murphy, E., et al. (2014). Relationship of DNA methylation and gene expression in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med, 190*, 1263-1272.

Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science, 356*, eaaj2239.

Yoo, S., Takikawa, S., Geraghty, P., Argmann, C., Campbell, J., Lin, L., et al. (2015). Integrative analysis of DNA methylation and gene expression data identifies EPAS1 as a key regulator of COPD. *PLoS Genet, 11*, e1004898.

Yousefi, P. D., Suderman, M., Langdon, R., Whitehurst, O., Davey Smith, G., & Relton, C. L. (2022). DNA methylation-based predictors of health: applications and statistical considerations. *Nat Rev Genet, 23*, 369-383.

Yu, M., Hon, G. C., Szulwach, K. E., Song, C. X., Zhang, L., Kim, A., et al. (2012). Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell, 149*, 1368-1380.

Yu, X., Zhao, H., Wang, R., Chen, Y., Ouyang, X., Li, W., et al. (2024). Cancer epigenetics: from laboratory studies and clinical trials to precision medicine. *Cell Death Discovery, 10*, 28.

Zeilinger, S., Kuhnel, B., Klopp, N., Baurecht, H., Kleinschmidt, A., Gieger, C., et al. (2013). Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One, 8*, e63812.

Zhang, X., Biagini Myers, J. M., Burleson, J. D., Ulm, A., Bryan, K. S., Chen, X., et al. (2018). Nasal DNA methylation is associated with childhood asthma. *Epigenomics, 10*, 629-641.

Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., et al. (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res, 38*, 4246-4253.

Zhao, S., Allis, C. D., & Wang, G. G. (2021). The language of chromatin modification in human cancers. *Nat Rev Cancer, 21*, 413-430.

Zhou, H., Liu, J., Zhou, C., Gao, N., Rao, Z., Li, H., et al. (2018). In vivo simultaneous transcriptional activation of multiple genes in the brain using CRISPR-dCas9-activator transgenic mice. *Nat Neurosci, 21*, 440-446.

Zhu, T., Liu, J., Beck, S., Pan, S., Capper, D., Lechner, M., et al. (2022). A pan-tissue DNA methylation atlas enables in silico decomposition of human tissue methylomes at cell-type resolution. *Nat Methods, 19*, 296-306.

Zhuang, B. C., Jude, M. S., Konwar, C., Ryan, C. P., Whitehead, J., Engelbrecht, H.-R., et al. (2024). Comparison of Infinium MethylationEPIC v2.0 to v1.0 for human population epigenetics:

considerations for addressing EPIC version differences in DNA methylation-based tools. *bioRxiv*, 2024.2007.2002.600461.

Zong, D., Liu, X., Li, J., Ouyang, R., & Chen, P. (2019). The role of cigarette smoke-induced epigenetic alterations in inflammation. *Epigenetics Chromatin, 12*, 65.