Introduction

COPD

Chronic obstructive pulmonary disease (COPD) is the most common respiratory disease, characterized by chronic and progressive course.[1] It's pathology involves chronic inflammatory response of the airways, overproduction of mucus (resulting in chronic bronchitis), parenchymal tissue destruction (resulting in emphysema) and abnormal repair defence mechanisms (resulting in small airway fibrosis).[2] This leads to air trapping in the lungs, sputum production, obstructed exhalation, dyspnoea and cough, common symptoms associated with COPD.[2] Although COPD can be stable over time, exacerbations, defined as an acute worsening of respiratory symptoms resulting in additional therapy, often occur.[3]

Epidemiology and risk factors

Chronic obstructive pulmonary disease is a major public health burden.[3,4] COPD is currently the third leading cause of death worldwide with more than 3 million deaths per year.[5,6] Although it is difficult to estimate the prevalence due to the variability in diagnostic criteria, recent standardized meta-analyses show a significant increase in both global and regional prevalence in 2010, compared with 1990.[7] In 2010, the global prevalence based on spirometry was estimated to be 11.7% with 384 million cases.[7] Prevalence is higher in current smokers and ex-smokers, in males compared with females and increases with age and air pollution.[3]

The COPD prevalence and annual deaths are predicted to increase, due to the increased prevalence of smoking and air pollution in some regions and aging of the population.[4] Exacerbations are an important reason for hospitalization and are responsible for about 10% of all acute medical admissions, adding to the mortality and morbidity rates and overall burden of the disease.[8] Survival rates of COPD patients with three or more exacerbations in 5 years follow-up are markedly reduced compared with those without exacerbations (30% versus 80%).[9]

Although smoking is a predominant risk factor, 25-45% of never-smokers also develop COPD.[10,11] It has been hypothesized that COPD is the result of a more complex interaction of cumulative exposures to noxious gases and particles (smoking, air pollution and/or occupational exposure) and a range of host factors, including (epi) genetic factors, poor lung growth, age and airway hyper-responsiveness.[3] From a genetic perspective, an important question to answer is to what extent the genetic determinants of COPD are overlapping in smokers and non-smokers or whether there are specific gene-environment interactions that change the genetic architecture in these two groups.

Diagnosis

According to the Global initiative for chronic Obstructive Lung Disease (GOLD) the COPD diagnosis is based on the airflow limitation, as measured by the lung function tests.[2] Spirometry is the most objective lung function test and the post-bronchodilator ratio of the forced expiratory volume in 1 second (FEV_1) over the forced vital capacity of the lungs (FVC) resulting in <0.7 is a standard definition of the airflow limitation.[3] However, using this fixed ratio results in more frequent over-diagnosis in the elderly (the lung function normally lowers with age), and more frequent under-diagnosis in younger adults (<45 years).[12] Thus, the American Thoracic Society (ATS) and the European Respiratory Society (ERS) guidelines recommend the lower limit of normal (LLN) as a cut-off value (FEV_1/FVC<LLN). LLN represents the lower 5% of the healthy population, evaluated by comparison with the reference values based on age, height, sex and race.[13] However, this value is highly dependent on the reference population. Since simplicity and consistency of a diagnostic tool are highly valued in clinical practice and research, GOLD still prefers the use of the fixed ratio[3] and is therefore widely used in genetic and epidemiological studies as well as in the studies described in this thesis. In the new assessment tool proposed by GOLD 2017,[3] COPD is classified in stages of severity based on the combination of severity of airflow limitation (FEV_1 % predicted), exacerbation history and symptoms burden (**Figure 1**).[2,3] GOLD is confident that this tool will result in a decrease of misclassification and better diagnosis and treatment of COPD.

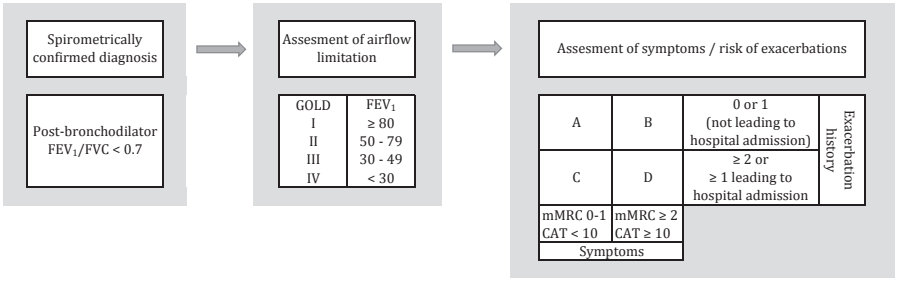

Figure 1. Combined COPD assessment tool proposed by GOLD 2017 (Adapted with permission from GOLD from "GOLD Management and Prevention of COPD 2017", Copyright © 2016 GOLD),[2] mMRC- Modified British Medical Research Council Questionnaire used for symptom assessment; CAT – COPD Assessment Test™.

Comorbidities

Various other pulmonary conditions are known to coexist with COPD and increase the severity of the disease. Those include asthma, pneumonia, pulmonary hypertension, pulmonary embolism, obstructive sleep apnoea, idiopathic pulmonary fibrosis and lung cancer.[14,15] Most are considered to be part of the COPD spectrum or a consequence of COPD pathology.[14] Asthma is considered to be a major risk factor for COPD, where people with asthma have 12-fold increased risk of COPD compared with those without asthma.[16] However, it is difficult to clinically differentiate asthma and COPD in adults as in 40% of the elderly it coexists with COPD.[17]

Furthermore, COPD is a systemic disorder that is associated with multiple extrapulmonary comorbid diseases.[18,19] Most common are cardiovascular diseases, metabolic diseases, cancer and depression, among many others.[15] The comorbidities may in part be explained by common factors such as smoking, alcohol, diet, ageing and polypharmacy or may share pathophysiological mechanisms and be consequence of the systemic inflammation.[15,18] Comorbidities have impact on the severity of the exacerbations and consequently on hospitalization rates and prognosis and are thus relevant for clinical care and management.[20] Depression is proposed to be one of the most underestimated, yet prevalent comorbidities of COPD[15] for which the common mechanisms are far from understood.[21] A total of 26% of COPD cases have depression, which has been associated with female gender, younger age, poor prognosis, smoking and severity of COPD with higher exacerbation risk.[22,23] Depression may be the result of (preclinical) pathology, which impacts quality of life. On the other hand, it has been speculated that there may be shared risk factors with effects on brain, such as smoking, ageing, hypoxaemia and systemic inflammation.[15,24] Alternatively, there may be shared genomics determinants.[15] In the present study, I studied the common genetic and epigenetic determinants of COPD, depression and other COPD related comorbidity.

OMICS OF COPD

The suffix *-omics* (from Greek word "ὁμοῖος" - common, general, one that concerns all parts) added to a molecular term denotes a comprehensive or global assessment of a set of molecules, which are collectively denoted with the suffix *–ome.*[25] Accordingly, *genomics, epigenomics, transcriptomics* and *metabolomics* represent a comprehensive study of a *genome, epigenome, transcriptome* and *metabolome,* respectively, the complete sets of different genes, transcripts of genes, proteins or active molecules (metabolites) of an organism (**Figure 2**).

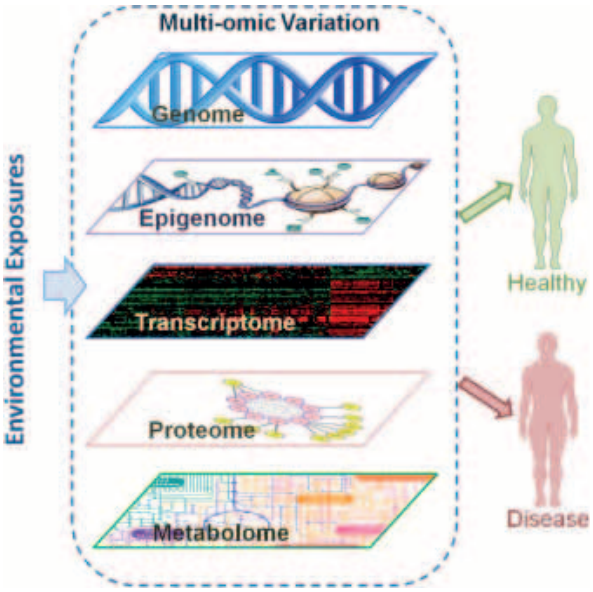

Figure 2. Multi-omics approach to studying a disease. Reprinted with permission from Elsevier. Sun YV, Hu YJ. Chapter Three-Integrative Analysis of Multi-omics Data for Discovery and Functional Studies of Complex Human Diseases. Advances in genetics. 2016 Dec 31;93:147-90. Copyright © 2016 Elsevier Inc.

Analyses that integrate these layers are powerful tools for understanding the pathogenesis and pathology of complex diseases.[25] Such integrative studies may improve our understanding of how specific genetic variations contribute to the disease.[26] The integration of data across multi-omics layers allows us to:

- gain understanding of the functional consequences and relevant interactions between different layers;[27]
- build pathways and networks based on a prior published or bioinformatic knowledge in order to understand the pathophysiology of a disease.[26,28]

There has been significant progress in understanding pulmonary diseases in recent years based on the development of omics research.[29] COPD is a complex disease with overlapping endophenotypes, which may be the result of interactions of many factors, both external and internal.[30] In this thesis I aim to disentangle the pathogenesis of COPD and its co-morbidity, using various omics approaches discussed below.

Genetics

Genetics focuses on identification of a DNA (Deoxyribonucleic acid) sequence changes, such as single nucleotide variations (SNVs). These may be associated with the risk and development of pathology, treatment response or prognosis.[26,31] The human genome is an important driver of the risk of COPD. The heritability of COPD is estimated to be 20-60%.[32,33] COPD as a complex disease is likely the result of the interplay of rare variants with moderate to large effects and common variants with small effects. Genetic studies identified several genetic risk factors for COPD. The first and most well-known genetic variant causing emphysema at young age is the rare variant in *SERPINA1* gene at chromosome 14q, resulting in Alpha-1-antitrypsin (AAT) deficiency.[34,35] Candidate-gene studies, focusing on genes encoding protein implicated in the pathogenesis of COPD, highlighted broad areas of the genome potentially involved in COPD, but did not yield informative reproducible results.[36] Genome-wide association studies (GWAS), using hypothesis-free and genome wide approach, have successfully identified common variants associated with COPD[37–43] and related outcomes, such as lung function measurements (FEV_1, FEV_1/FVC),[37,44–47] emphysema,[48] chronic bronchitis.[41,49] Findings are not only replicable within an endophenotype, but also show a substantial overlap accross.[43] The loci identified in COPD GWASs that were replicated include Hedgehog-interacting protein (*HHIP*), Family with sequence similarity 13 member A (*FAM13A*), Nicotinic cholinergic receptors (*CHRNA3/5*), Ion-responsive element binding protein 2 (*IREB2*), Cytochrome P450 family gene (*CYP2A6*), Member RAS oncogene family gene (*RAB4B*) and Egl-9 family hypoxic-inducible factor 2 (*EGLN2*).[37,43]

As has been the case in many other disorders, the use of endophenotypes, i.e., continuous heritable traits that are associated with the disease (diagnosis), has been even more successful in identifying genetic loci.[50] The major advantage of this approach is that it overcomes the problems of diagnostic classification, which for many disorders including COPD is arbitrary and may introduce misclassification. The use of endophenotypes results in loss of specificity as there is no 1:1 relationship between the endophenotype and the disease and endophenotype may be related to multiple disorders.[51] Yet, there is a gain in efficiency because the endophenotypes often have a higher heritability than the disease and are usually available in large number of persons, covering a full range of disease severity: from healthy, pre-clinic, moderate to severe. Based on a genome-wide association discovery in 48,943 individuals and follow-up in 95,375 individuals, Wain et al. reported 97 loci relevant for lung function, of which 43 were novel.[37] **Figure 3** gives an overview of the 97 loci, underlying those relevant for COPD.

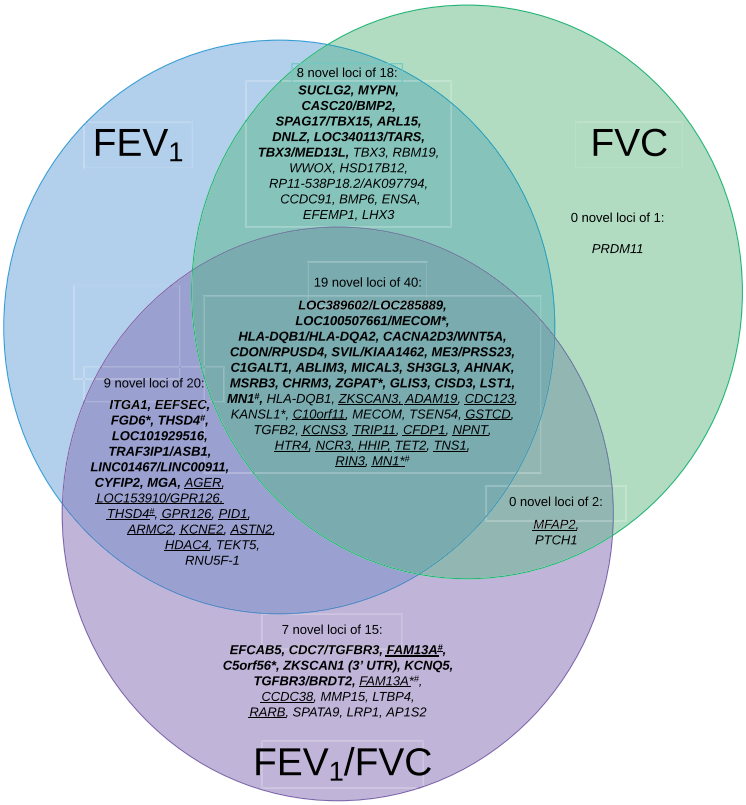

Figure 3. Loci associated with lung function related to COPD.[50] In bold - novel findings. Underlined – loci associated with COPD ($P<5.26\times10^{-4}$). *Loci associated with smoking. #Same gene has 2 variants one novel, one already known.

The genetic risk score derived from these is associated with COPD susceptibility results in 3.7-fold difference in COPD risk between highest and lowest genetic risk score deciles (**Figure 4**).[37] The odds ratios per standard deviation of the risk score (~6 alleles) (95% confidence interval) is 1.24 (1.20-1.27), $P=5.05\times10^{-49}$ show a consistent increase over the full distribution.

When interpreting the biological and physiological pathways the 97 genetic variants are implicated in those involved in development, elastic fibres and epigenetic regulation pathways. These pathways point to targets for drugs and compounds in development for COPD and asthma.

Despite the successes, a large part of the estimated heritability is still missing. This may be explained by:

- Rare variants that are not well covered to date by GWASs;
- Gene interactions;
- Epigenetic modifications that are in part driven by genetic variants.

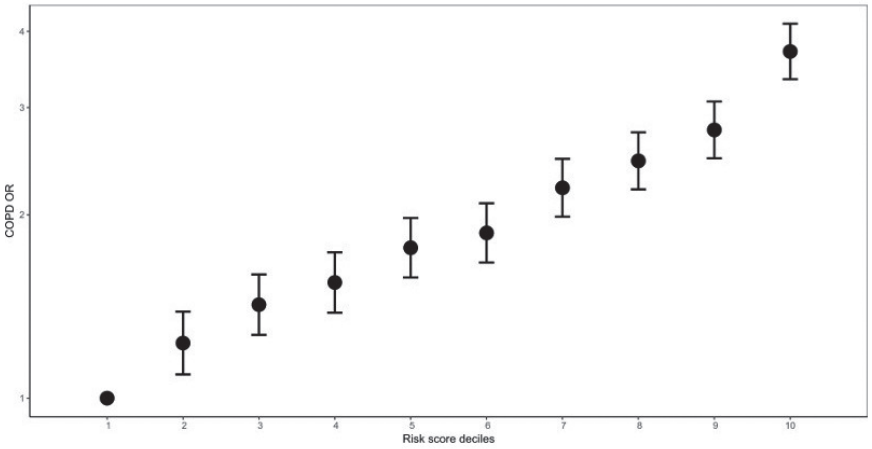

Figure 4. Odds ratios for spirometrically-defined COPD for weighted genetic risk score deciles in UK Biobank (10,547 cases, pre-bronchodilator % predicted FEV_1<80% and FEV_1/FVC<0.7, and 53,948 controls, FEV_1/FVC>0.7 and % predicted FEV_1>80%, weights derived from non-discovery populations). For each decile, odds ratios were obtained using a logistic regression adjusted for age, age2, sex, height, smoking status, pack-years and the first 10 ancestry principal components. Source: Wain et al.[37]

So far GWAS has brought to surface common variants. Rare variants are not covered by the arrays used for GWAS, but, more importantly, are also not well imputed using common reference panels (e.g. HapMap and 1000 Genomes).[52,53] Of note is that imputation is improving with larger reference panels, such as the Haplotype Reference Consortium panel (HRC) combining several widely used panels (with total of 64,976 haplotypes) and data from exome sequencing. Using HRC, rare variants can be imputed more reliably in GWAS.[52] An alternative route to discover rare variants is family based studies. While a variant is rare in the general population, within a family of first- and second-degree relatives such variant will be transmitted with a 50% probability. Thus, within a family, the variant is common. To find rare variants Qiao et al conducted a whole exome sequencing analysis in 2,543 subjects from two family-based studies.[54] Applying a gene-based segregation test in the family-based data, they identified significant segrega-

tion of rare loss of function variants in *TBC1D10A* and *RFPL1* ($P<2\times10^{-6}$) but were unable to find similar variants in the case–control study. Further, they identified individuals with putative high-risk variants, including patients harbouring homozygous mutations in genes associated with cutis laxa and Niemann–Pick Disease Type C.[54] Also a recent whole genome sequencing study in severe COPD identified a large number of potentially important functional variants, with the strongest associations being in known COPD risk loci, including *HHIP* and *SERPINA1*.[55] Encouraged by these findings, in this thesis I also used a family-based approach to identify rare variants implicated in COPD.

Epigenomics

Epigenomics investigates epigenome which is a set of chemical modifications of the chromatin and DNA molecule that regulate gene expression, without changing the DNA sequence.[26,28] These changes are usually reversible, and may be driven by genetic (heritable) and environmental factors. Of note is that in some instances the modifications may be permanent,[56,57] and cell-type (tissue) specific.[58]

The most commonly studied epigenetic mechanisms are DNA methylation and histone modifications.[29] DNA methylation is addition of a methyl group ($-CH_3$) to any cytosine (C) that is next to guanine (G) in the DNA sequence, converting it to 5-methylcytosine. These sites are called CpGs (short for 5'-C-phosphate-G-3'), and in humans around 70-80% of CpGs are methylated.[59] Epigenome-wide association studies (EWAS) have shown that differential DNA methylation patterns have a role in the disease development.[60] It has also been shown that smoking affects DNA methylation,[56,61,62] which in turn may lead to the disease. Furthermore, genetic variants may modulate regulatory mechanisms such as DNA methylation (methylation quantitative trait loci - meQTLs).[63] Epigenetic studies of COPD have identified differential DNA methylation associated with COPD severity, poor lung function and use of systemic corticosteroids.[64–66] It has been postulated that early exposure to risk factors, such as maternal tobacco smoking during pregnancy, are associated with risk of asthma and lower lung function, through changes in DNA methylation.[67] This may also affect the risk of COPD at old age. When combining epigenome and transcriptome data from lung tissues of COPD patients and controls, *EPAS1* gene has been proposed as a key regulator of COPD pathogenesis and has been confirmed by functional studies, highlighting the need for integrative studies.[68] This gene has not emerged in the list of genes implicated in COPD or endophenotypes to date.

In this thesis, I addressed the specific question whether the GWAS variants change the epigenome landscape and subsequently alter the transcription of the gene, integrating genetic, epigenetic and transcriptomic data. GWAS has been extremely successful, but the functional effects of the identified genes in COPD pathogenesis

were largely not investigated. Another poorly understood issue is the interaction of the genetic drivers of pathology with the environment. Integrating genetic research with other -omics may improve our understanding of functional effects and gene interactions, since at the omics level such effects are expected to be larger than at the level of a complex disease such as COPD, which involves a large range of phenotypes and comorbidities driven by both external and internal factors.[30] In this thesis, I aimed to understand the functional changes driving the association of GWAS hits to COPD at the level of epigenomics and transcriptomics. I further use genetics to address the question whether a common genetic background explains the comorbidity in COPD occurring in patients.

Transcriptomics

Transcriptomics explores genome-wide levels of RNA transcripts (gene expression) both qualitatively and quantitatively, which are directly influenced by the genome (expression quantitative trait loci – eQTLs) and epigenome (expression quantitative trait methylation – eQTMs),[26] besides environmental factors. It is known that gene expression can be tissue specific and in order to investigate a disease one should focus on the tissue of interest. One study showed that environmental risk factors such as smoking influences the transcriptome of the small airway epithelium,[69] even after smoking cessation.[70] However, some genes are expressed globally over tissues. An important issue to consider is that multiple tissues may be involved in a disease. Smoking, the major determinant of COPD, may affect the expression in blood, lung tissue or other tissues. Indeed, a study investigating blood of smokers with and without COPD, could discriminate the cases from the controls based on the expression profile of 26 genes involved in immune and inflammatory response and sphingolipid metabolism.[71] Although transcriptomic studies were useful in identifying specific gene expression pattern associated with COPD[72,73] and with drug response,[74,75] a global expression profile unique for COPD has not been found.[29] In this thesis, I chose to integrate genomics with gene expression to explore the functional effects of genetic and epigenetic changes.

Metabolomics

Metabolomic studies all metabolites present in a tissue, which are small molecules (<1 kDa) of endogenous or exogenous etiology.[29] These include peptides, amino acids, nucleic acids, carbohydrates, vitamins, polyphenols, and alkaloids, among other compounds that are involved in cellular metabolic functions. In pulmonary research of metabolomics, studied samples include blood, sputum, exhaled breath condensate, bronchoalveolar lavage fluid and lung tissue.[76] The identification

of changes in biomarkers that can identify or differentiate various disease phenotypes even in the early stages is of high importance in COPD.[76] Several studies used metabolomics methods to investigate biochemical effects induced by COPD, exacerbations and its related outcomes as well as external effects of smoking and drugs, using different samples.[76] Most of the studies identified metabolites involved in systemic inflammation, protein degradation and oxidative stress.[77–79] Consistent with the transcriptomics studies in blood, mentioned above, another study of lipids in sputum reported that sphingolipids were highly expressed in sputum of smokers with COPD compared with smoking controls.[80,81] However, these studies were very limited in sample size, therefore the results should be further confirmed in larger samples. In this thesis I have combined the data of two large population-based studies to understand the metabolomics changes in COPD. As a person's metabolism may change causing the disease or change as a result of the disease process, I used a genomic method, explained below, to disentangle these effects.

Mendelian Randomization

A major problem in observational epidemiological studies and the translation of findings to the clinics is the problem of causal inferences due to the possible reverse causation: e.g. to distinguish whether the metabolic or other omics changes are causing a disease or are the consequence of the pathology. One of the most important approaches developed in the omics era is the method referred to as Mendelian Randomization (MR). MR is a cross-omics approach, which uses genetic data as an instrumental variable (IV) to examine the evidence for causal effects between modifiable exposures (risk factors) and an outcome (disease).[82] The rationale is that similar to randomized controlled trials, the genotypes are assigned randomly and the disease starts after meiosis.[83] Randomisation is based on Mendel's second law that the inheritance of one trait is independent of the inheritance of other traits.[83] The IV (usually based on a combination of genotypes that are associated to the disease) has to comply with three assumptions: (1) to be associated with the exposure; (2) to be independent of any confounders of the exposure-outcome association and (3) to be related to the outcome only through the exposure.[83] MR analysis can be conducted unilateral, testing a specific hypothesis, e.g. if alcohol consumption is casually related to the risk of cardiovascular mortality.[84] In the setting of multi- or cross-omics research as in the metabolomics-COPD study I performed, the MR is often bi-directional, testing the hypothesis that: 1) the metabolite is causally related to COPD and therefore the genetic determinants of metabolite (used as instrumental variable) are also associated to COPD and 2) (pre)clinical COPD pathology affect the metabolite levels, which translates into the model where genes determining COPD are also associated to metabolite.

SCOPE OF THIS THESIS

The overall aim of this thesis is to identify novel molecular determinants of COPD, lower lung function and related pathology such as depression and to perform integrative studies to investigate the functional role and interaction of multiple omics layers.

In **Chapter 2** I investigate COPD applying different omics approaches. In **Chapter 2.1,** I describe a genome-wide linkage scan performed in a search for rare genetic variants which have a role in familial COPD, utilizing family-based settings of the Erasmus Rucphen Family (ERF) study and integrating the data from the Rotterdam Study (RS), the LifeLines study (LLS), Hobbs et al.[85] and the Vlagtwedde/Vlaardingen study. **Chapter 2.2** and **Chapter 2.3** investigate the functional role of two established COPD GWAS loci by exploring a multi-omics approach linking the genetic loci to the epigenomic and transcriptomic effects in the Rotterdam study and the Lung expression quantitative loci mapping study. **Chapter 2.2** examines the chromosome 15q25 locus and its meQTL effects in blood and eQTL effects in lung tissue, to understand the functional effects of this locus in relation to COPD. Similarly, **Chapter 2.3** investigates a top variant from a novel locus on 19q13, identified in COPD GWAS, and mediation of its genetic risk on gene expression, through DNA methylation signatures. In **Chapter 2.4,** I present an EWAS meta-analysis of lung function levels in never-smokers only, to identify factors other than smoking which affect lung function through DNA methylation in RS and LLS.

In **Chapter 3**, the thesis focuses on comorbidities of COPD, including early and late metabolic effects. **Chapter 3.1** describes a large meta-analysis in Pregnancy And Childhood Epigenetics (PACE) consortium studying DNA methylation in relation to lung function at birth and the effects on lung function, asthma and COPD throughout life course. In **Chapter 3.2,** I study circulating metabolites in relation with COPD in ERF, RS and several replication cohorts and apply multi-omics Mendelian Randomization approach to investigate causal relations of the metabolite-COPD associations. In **Chapter 3.3**, I use an integrative genetic approach to overlap genetic drivers of COPD and its non-pulmonary comorbidity. In **Chapter 3.4,** I investigate DNA methylation patterns specific for depression in a largest to date EWAS study in Cohorts for Heart and Aging in Genomic Epidemiology (CHARGE) consortium with the view to determine the overlap with that seen in COPD.

The main findings and implications described in my thesis I discuss in the **Chapter 4,** which I summarize in English and in Dutch in **Chapter 5.**

REFERENCES

1. European Respiratory Society. European lung white book. http://www.erswhitebook.org/chapters/chronic-obstructive-pulmonary-disease/. Published 2016.
2. Vestbo J, Hurd SS, Agustí AG, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease GOLD executive summary. *Am J Respir Crit Care Med*. 2013;187(4):347-365. doi:10.1164/rccm.201204-0596PP
3. GOLD. From the Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease. *Available from http// goldcopd org*. 2017. http://goldcopd.org.
4. World Health Organization. Projections of global mortality and burden of disease from 2015 to 2030. *Heal Stat Inf Syst*. 2013;3(11):2015-2030. doi:978 92 4 156422 9
5. Naghavi M, Wang H, Lozano R, et al. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2015;385(9963):117-171. doi:10.1016/S0140-6736(14)61682-2
6. Lozano R, Naghavi M FK. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380(380):2095–2128. doi:https://doi.org/10.1016/S0140-6736(12)61728-0
7. Adeloye D, Chua S, Lee C, et al. Global and regional estimates of COPD prevalence: Systematic review and meta–analysis. *J Glob Health*. 2015;5(2). doi:10.7189/jogh.05.020415
8. Roberts CM, Stone RA, Lowe D, Pursey NA, Buckingham RJ. Co-morbidities and 90-day outcomes in hospitalized COPD exacerbations. *COPD J Chronic Obstr Pulm Dis*. 2011;8(5):354-361. doi:10.3109/15412555.2011.600362
9. Soler-Cataluna JJ. Severe acute exacerbations and mortality in patients with chronic obstructive pulmonary disease. *Thorax*. 2005;60(11):925-931. doi:10.1136/thx.2005.040527
10. Lamprecht B, McBurnie MA, Vollmer WM, et al. COPD in Never Smokers\nResults From the Population-Based Burden of Obstructive Lung Disease Study. *Chest*. 2011;139(April):752-763. doi:10.1378/chest.10-1253
11. Pavord ID, Yousaf N, Birring SS, Salvi SS, Barnes PJ. Chronic obstructive pulmonary disease in non-smokers. *Lancet*. 2009;374(9691):1964; author reply 1965--1966. doi:10.1016/S0140-6736(09)62114-0
12. Celli BR, Macnee W. Standards for the diagnosis and treatment of patients with COPD : a summary of the ATS / ERS position paper. *Eur Respir J*. 2004;23(6):932-946. doi:10.1183/09031936.04.00014304
13. van Dijk W, Tan W, Li P, Best G, Li S. Clinical Relevance of Fixed Ratio vs Lower Limit of Normal of FEV_1:FVC in COPD: Patient-Reported Out.pdf. *Ann Fam Med*. 2015;13(1):41-48.
14. Chatila WM, Thomashow BM, Minai OA, Criner GJ, Make BJ. Comorbidities in Chronic Obstructive Pulmonary Disease. *Proc Am Thorac Soc*. 2008;5(4):549-555. doi:10.1513/pats.200709-148ET
15. Barnes PJ, Celli BR. Systemic manifestations and comorbidities. *Eur Respir J*. 2009;33(5):1165-1185. doi:10.1183/09031936.00128008
16. Silva GE. Asthma as a Risk Factor for COPD in a Longitudinal Study. *Chest*. 2004;126(1):59-65. doi:10.1378/chest.126.1.59

17. Aryal S, Diaz-guzman E, Mannino DM. Asthma Treatment Options in Asthma and Chronic Obstructive Pulmonary Diseases Overlap Syndrome. *Touch Briefings, Eur Respir Dis*. 2011; 7(2):101-105.
18. Divo M, Cote CG, De Torres JP, et al. Comorbidities and risk of mortality in patients with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2012;186(2):155-161. doi: 10.1164/rccm.201201-0034OC
19. Almagro P, Cabrera FJ, Diez-Manglano J, et al. Comorbidome and short-term prognosis in hospitalised COPD patients: The ESMI study. *Eur Respir J*. 2015;46(3):850-853. doi: 10.1183/09031936.00008015
20. Mannino DM, Thorn D, Swensen A, Holguin F. Prevalence and outcomes of diabetes, hypertension and cardiovascular disease in COPD. *Eur Respir J*. 2008;32(4):962-969. doi: 10.1183/09031936.00012408
21. Norwood RJ. A review of etiologies of depression in COPD. *Int J COPD*. 2007;2(4):485-491.
22. Hanania NA, Müllerova H, Locantore NW, et al. Determinants of depression in the ECLIPSE chronic obstructive pulmonary disease cohort. *Am J Respir Crit Care Med*. 2011;183(5):604-611. doi:10.1164/rccm.201003-0472OC
23. Maurer J, Rebbapragada V, Borson S, et al. Anxiety and depression in COPD: Current understanding, unanswered questions, and research needs. *Chest J*. 2008;134(4 SUPPL.):43S-56S. doi:10.1378/chest.08-0342
24. Anisman H, Merali Z, Hayley S. Neurotransmitter, peptide and cytokine processes in relation to depressive disorder: Comorbidity between depression and neurodegenerative disorders. *Prog Neurobiol*. 2008;85(1):1-74. doi:10.1016/j.pneurobio.2008.01.004
25. Wijmenga C, Zhernakova A. The importance of cohort studies in the post-GWAS era. *Nat Genet*. 2018;50(3):322-328. doi:10.1038/s41588-018-0066-3
26. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. *Genome Biol*. 2017;18(1):83. doi:10.1186/s13059-017-1215-1
27. Valdes AM, Glass D, Spector TD. Omics technologies and the study of human ageing. *Nat Rev Genet*. 2013;14(9):601-607. doi:10.1038/nrg3553
28. Sun Y V., Hu Y-J. Integrative Analysis of Multi-omics Data for Discovery and Functional Studies of Complex Human Diseases. *Adv Genet*. 2016;93:147-190. doi:10.1016/bs.adgen.2015.11.004
29. Kan M, Shumyatcher M, Himes BE. Using omics approaches to understand pulmonary diseases. *Respir Res*. 2017;18(1):149. doi:10.1186/s12931-017-0631-9
30. Huang Q. Genetic Study of Complex Diseases in the Post-GWAS Era. *J Genet Genomics*. 2015; 42(3):87-98. doi:10.1016/j.jgg.2015.02.001
31. Manichaikul A, Nguyen JN. Genetic studies as a tool for identifying novel potential targets for treatment of COPD. *Eur Respir J*. 2017;50(5):1702042. doi:10.1183/13993003.02042-2017
32. Ingebrigtsen T, Thomsen SF, Vestbo J, et al. Genetic influences on Chronic Obstructive Pulmonary Disease. *a Twin Study*. 2010;104(12):1890-1895. doi:10.1016/j.rmed.2010.05.004
33. Zhou JJ, Cho MH, Castaldi PJ, Hersh CP, Silverman EK, Laird NM. Heritability of Chronic Obstructive Pulmonary Disease and Related Phenotypes in Smokers. *Am J Respir Crit Care Med*. 2013;188(8):941-947. doi:10.1164/rccm.201302-0263OC [doi]
34. Bashir, A., Shah, N.N., Hazari, Y.M., Habib, M., Bashir, S., Hilal, N., Banday, M., Asrafuzzaman, S., and Fazili KM. Novel variants of SERPIN1A gene: Interplay between alpha1- antitrypsin deficiency and chronic obstructive pulmonary disease. *Respir Med*. 2016;117:139-149. doi: https://doi.org/10.1016/j.rmed.2016.06.005

35. Laurell CB ES, Laurell CB ES. The electrophoretic alpha 1-globulin pattern of serum in alpha 1-antitrypsin deficiency. *Scand J Clin Lab Invest*. 1963;15(2):132-140. doi:https://doi.org/10.1080/00365516309051324
36. Hobbs BD, Hersh CP. Integrative genomics of chronic obstructive pulmonary disease. *Biochem Biophys Res Commun*. 2014;452(2):276-286. doi:10.1016/J.BBRC.2014.07.086
37. Wain L V., Shrine N, Artigas MS, et al. Supplementary: Genome-wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. *Nat Genet*. 2017;49(3):416-425. doi:10.1038/ng.3787
38. Kim DK, Cho MH, Hersh CP, et al. Genome-wide association analysis of blood biomarkers in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2012;186(12):1238-1247. doi:10.1164/rccm.201206-1013OC
39. Boezen HM. Genome-wide association studies: what do they teach us about asthma and chronic obstructive pulmonary disease? *Proc Am Thorac Soc*. 2009;6(8):701-703. doi:10.1513/pats.200907-058DP
40. Pillai SG, Ge D, Zhu G, et al. A Genome-Wide association study in chronic obstructive pulmonary disease (COPD): Identification of two major susceptibility loci. *PLoS Genet*. 2009;5(3):e1000421. doi:10.1371/journal.pgen.1000421
41. Silverman EK, Mosley JD, Palmer LJ, et al. Genome-wide linkage analysis of severe , early-onset chronic obstructive pulmonary disease : airflow obstruction and chronic bronchitis phenotypes. *Hum Mol Genet*. 2002;11(6):623-632. http://dx.doi.org/10.1093/hmg/11.6.623.
42. Cho MH, McDonald M-LN, Zhou X, et al. Risk loci for chronic obstructive pulmonary disease: a genome-wide association study and meta-analysis. *Lancet Respir Med*. 2014;2(3):214-225. doi:10.1016/S2213-2600(14)70002-5
43. Hobbs BD, De Jong K, Lamontagne M, et al. Genetic loci associated with chronic obstructive pulmonary disease overlap with loci for lung function and pulmonary fibrosis. *Nat Genet*. 2017;49(3):426-432. doi:10.1038/ng.3752
44. van der Plaat DA, de Jong K, Lahousse L, et al. Genome-wide association study on the FEV_1/FVC ratio in never-smokers identifies HHIP and FAM13A. *J Allergy Clin Immunol*. 2017;139(2):533-540. doi:10.1016/j.jaci.2016.06.062
45. Hansel NN, Ruczinski I, Rafaels N, et al. Genome-wide study identifies two loci associated with lung function decline in mild to moderate COPD. *Hum Genet*. 2013;132(1):79-90. doi:10.1007/s00439-012-1219-6
46. Repapi E, Sayers I, Wain L V., et al. Genome-wide association study identifies five loci associated with lung function. *Nat Genet*. 2010;42(1):36-44. doi:10.1038/ng.501
47. Artigas MS, Loth DW, Wain L V., et al. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nat Genet*. 2011;43(11):1082-1090. doi:10.1038/ng.941
48. Castaldi PJ, Cho MH, Estépar RSJ, et al. Genome-wide association identifies regulatory loci associated with distinct local histogram emphysema patterns. *Am J Respir Crit Care Med*. 2014;190(4):399-409. doi:10.1164/rccm.201403-0569OC
49. Lee JH, Cho MH, Hersh CP, et al. Genetic susceptibility for chronic bronchitis in chronic obstructive pulmonary disease. *Respir Res*. 2014;15(1):113. doi:10.1186/s12931-014-0113-2
50. Beauchaine TP, Constantino JN. Redefining the endophenotype concept to accommodate transdiagnostic vulnerabilities and etiological complexity. *Biomark Med*. 2017;11(9):769. doi:10.2217/BMM-2017-0002

51. Hasler G, Drevets WC, Manji HK, Charney DS. Discovering Endophenotypes for Major Depression. *Neuropsychopharmacology*. 2004;29(10):1765-1781. doi:10.1038/sj.npp.1300506
52. Iglesias AI, van der Lee SJ, Bonnemaijer PWM, et al. Haplotype reference consortium panel: Practical implications of imputations with large reference panels. *Hum Mutat*. 2017;38(8): 1025-1032. doi:10.1002/humu.23247
53. Huang J, Howie B, McCarthy S, et al. Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. *Nat Commun*. 2015;6:8111. doi:10.1038/ncomms9111
54. Qiao D, Ameli A, Prokopenko D, et al. Whole exome sequencing analysis in severe chronic obstructive pulmonary disease. *Hum Mol Genet*. 2018;27(21):3801-3812. doi:10.1093/hmg/ddy269
55. Prokopenko D, Sakornsakolpat P, Fier HL, et al. Whole-Genome Sequencing in Severe Chronic Obstructive Pulmonary Disease. *Am J Respir Cell Mol Biol*. 2018;59(5):614-622. doi: 10.1165/rcmb.2018-0088OC
56. Joehanes R, Just AC, Marioni RE, et al. Epigenetic Signatures of Cigarette Smoking. *Circ Cardiovasc Genet*. 2016;9(5):436-447. doi:10.1161/CIRCGENETICS.116.001506
57. Ladd-Acosta C, Fallin MD. The role of epigenetics in genetic and environmental epidemiology. *Epigenomics*. 2016;8(2):271-283. doi:10.2217/epi.15.102
58. Taudt A, Colomé-Tatché M, Johannes F. Genetic sources of population epigenomic variation. *Nat Rev Genet*. 2016;17(6):319-332. doi:10.1038/nrg.2016.45
59. Jabbari K, Bernardi G. Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. *Gene*. 2004;333(SUPPL.):143-149. doi:10.1016/j.gene.2004.02.043
60. M.B. T, L. D-C, N. V-R, H.C. W. DNA methylation in white blood cells: Association with risk factors in epidemiologic studies. *Epigenetics*. 2011;6(7):828-837. doi:http://dx.doi.org/10.4161/epi.6.7.16500
61. Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-smoking-related differential DNA methylation: 27 K discovery and replication. *Am J Hum Genet*. 2011;88(4):450-457. doi: 10.1016/j.ajhg.2011.03.003
62. Wan ES, Qiu W, Carey VJ, et al. Smoking-associated site-specific differential methylation in buccal mucosa in the COPDGene study. *Am J Respir Cell Mol Biol*. 2015;53(2):246-254. doi: 10.1165/rcmb.2014-0103OC
63. Shi J, Marconett C, Duan J, Hyland P, … PL-N, 2014 undefined. Characterizing the genetic basis of methylome diversity in histologically normal human lung tissue. *NatureCom*. 2014. https://www.nature.com/articles/ncomms4365.
64. Qiu W, Baccarelli A, Carey VJ, et al. Variable DNA methylation is associated with chronic obstructive pulmonary disease and lung function. *Am J Respir Crit Care Med*. 2012;185(4): 373-381. doi:10.1164/rccm.201108-1382OC
65. Wan ES, Qiu W, Baccarelli A, et al. Systemic steroid exposure is associated with differential methylation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2012; 186(12):1248-1255. doi:10.1164/rccm.201207-1280OC
66. Morrow JD, Cho MH, Hersh CP, et al. DNA methylation profiling in human lung tissue identifies genes associated with COPD. *Epigenetics*. 2016;11(10):730-739. doi: 10.1080/15592294.2016.1226451
67. Holloway JW, Meta-analysis GC, Joubert BR, et al. DNA Methylation in Newborns and Maternal Smoking in Pregnancy : Genome-wide Consortium Meta-analysis ARTICLE DNA

Methylation in Newborns and Maternal Smoking in Pregnancy : *Am J Hum Genet*. 2016; 98(April):680-696. doi:10.1016/j.ajhg.2016.02.019

68. Yoo S, Takikawa S, Geraghty P, et al. Integrative Analysis of DNA Methylation and Gene Expression Data Identifies EPAS1 as a Key Regulator of COPD. *PLoS Genet*. 2015;11(1):e1004898. doi:10.1371/journal.pgen.1004898
69. Hackett NR, Butler MW, Shaykhiev R, et al. RNA-Seq quantification of the human small airway epithelium transcriptome. *BMC Genomics*. 2012;13(1):82. doi:10.1186/1471-2164-13-82
70. Beane J, Sebastiani P, Liu G, Brody JS, Lenburg ME, Spira A. Reversible and permanent effects of tobacco smoke exposure on airway epithelial gene expression. *Genome Biol*. 2007;8(9): R201. doi:10.1186/gb-2007-8-9-r201
71. Bahr TM, Hughes GJ, Armstrong M, et al. Peripheral blood mononuclear cell gene expression in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol*. 2013;49(2):316-323. doi: 10.1165/rcmb.2012-0230OC
72. Singh D, Fox SM, Tal-Singer R, et al. Induced sputum genes associated with spirometric and radiological disease severity in COPD ex-smokers. *Thorax*. 2011;66(6):489-495. doi:10.1136/thx.2010.153767
73. Almansa R, Socias L, Sanchez-Garcia M, et al. Critical COPD respiratory illness is linked to increased transcriptomic activity of neutrophil proteases genes. *BMC Res Notes*. 2012;5:401. doi:10.1186/1756-0500-5-401
74. Masuno K, Haldar SM, Jeyaraj D, et al. Expression profiling identifies klf15 as a glucocorticoid target that regulates airway hyperresponsiveness. *Am J Respir Cell Mol Biol*. 2011;45(3):642-649. doi:10.1165/rcmb.2010-0369OC
75. Misior AM, Deshpande DA, Loza MJ, Pascual RM, Hipp JD, Penn RB. Glucocorticoid- and protein kinase A-dependent transcriptome regulation in airway smooth muscle. *Am J Respir Cell Mol Biol*. 2009;41(1):24-39. doi:10.1165/rcmb.2008-0266OC
76. Nobakht M Gh BF, Aliannejad R, Rezaei-Tavirani M, Taheri S, Oskouie AA. The metabolomics of airway diseases, including COPD, asthma and cystic fibrosis. *Biomarkers Biochem Indic Expo response, susceptibility to Chem*. 2015;20(1):5-16. doi:10.3109/1354750X.2014.983167
77. Chen Q, Deeb RS, Ma Y, Staudt MR, Crystal RG, Gross SS. Serum metabolite biomarkers discriminate healthy smokers from COPD smokers. *PLoS One*. 2015;10(12):e0143937. doi:10.1371/journal.pone.0143937
78. Adamko DJ, Nair P, Mayers I, Tsuyuki RT, Regush S, Rowe BH. Metabolomic profiling of asthma and chronic obstructive pulmonary disease: A pilot study differentiating diseases. *J Allergy Clin Immunol*. 2015;136(3):571-580.e3. doi:10.1016/j.jaci.2015.05.022
79. Ubhi BK, Riley JH, Shaw PA, et al. Metabolic profiling detects biomarkers of protein degradation in COPD patients. *Eur Respir J*. 2012;40(2):345-355. doi:10.1183/09031936.00112411
80. Bowler RP, Jacobson S, Cruickshank C, et al. Plasma Sphingolipids Associated with Chronic Obstructive Pulmonary Disease Phenotypes. *Am J Respir Crit Care Med*. 2015;191(3):275-284. doi:10.1164/rccm.201410-1771OC
81. Telenga ED, Hoffmann RF, T'Kindt R, et al. Untargeted lipidomic analysis in chronic obstructive pulmonary disease uncovering sphingolipids. *Am J Respir Crit Care Med*. 2014;190(2): 155-164. doi:10.1164/rccm.201312-2210OC
82. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?*. *Int J Epidemiol*. 2003;32(1):1-22. doi:10.1093/ije/dyg070

83. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008; 27(8):1133-1163. doi:10.1002/sim.3034
84. Millwood IY, Walters RG, Mei XW, et al. Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China. *Lancet (London, England)*. 2019;393(10183):1831-1842. doi:10.1016/S0140-6736(18)31772-0
85. Hobbs BD, Parker MM, Chen H, et al. Exome array analysis identifies a common Variant in IL27 associated with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2016; 194(1):48-57. doi:10.1164/rccm.201510-2053OC