Universidad de La Laguna

Escuela de Doctorado y Estudios de Posgrado

TÍTULO DE LA TESIS DOCTORAL

Uncovering the genetic variation involved in asthma exacerbations through multiple genomic approaches

AUTOR/A

Esther	Herrera	Luis

DIRECTOR/A

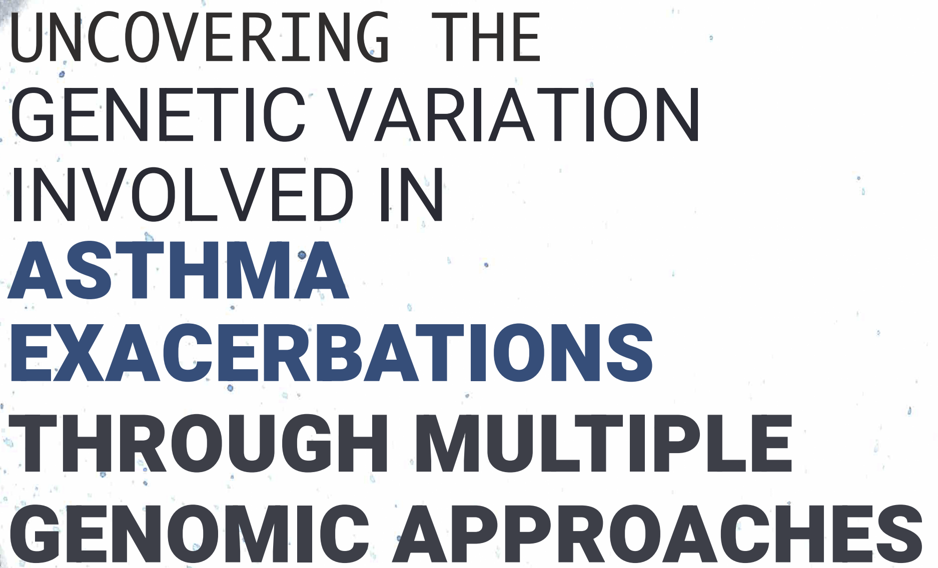
María del Mar	del Pino	Yanes

CODIRECTOR/A

DEPARTAMENTO O INSTITUTO UNIVERSITARIO

FECHA DE LECTURA

29/04/22



UNCOVERING THE GENETIC VARIATION INVOLVED IN **ASTHMA EXACERBATIONS** THROUGH MULTIPLE GENOMIC APPROACHES

DOCTORAL THESIS

Esther Herrera Luis

Universidad de La Laguna

PROGRAMA DOCTORADO CIENCIAS DE LA SALUD

Verificado ANECA, Resol. 19-3-2014

Uncovering the genetic variation involved in asthma exacerbations through multiple genomic approaches

Esther Herrera Luis

Supervisor: Dr. María del Mar del Pino Yanes

Department of Biochemistry, Microbiology, Cell Biology, and Genetics
University of La Laguna

This dissertation is submitted for the degree of
Ph.D. in Health Sciences

May 2022

Cover illustration: Tamara Ruiz Kolb

Printing: Acosta Gráfica S.L. / Print And Copy

Financial support for the printing of this doctoral thesis was provided by the Spanish Ministry of Science and Innovation MCIN/AEI/10.13039/501100011033 (grant PID2020-116274RB-I00).

Compendium statement

Dr. María del Mar del Pino Yanes, director of the doctoral thesis presented by the graduated **Ms. Esther Herrera Luis**

Certifies that

The doctoral thesis entitled *"Uncovering the genetic variation involved in asthma exacerbations through multiple genomic approaches"*, elaborated by Ms. Esther Herrera Luis under my supervision at the Department of Biochemistry, Microbiology, Cell Biology, and Genetics of University of La Laguna, meets the conditions of quality and scientific rigor to be presented and defended in the form of a compendium before the commission appointed for this purpose to qualify for the degree of Doctor of Health Sciences with the International Doctorate mention at University of La Laguna.

San Cristobal de La Laguna, 28 th March, 2022

Dr. María del Mar del Pino Yanes

Funding

This doctoral thesis was funded by the Spanish Ministry of Science and Innovation MCIN/AEI/10.13039/501100011033 and the European Regional Development Fund "ERDF A way of making Europe" by the European Union (grant SAF2017-83417R), and by the MCIN/AEI/10.13039/501100011033 (grant PID2020-116274RB-I00). My fellowship for the Predoctoral Research Personnel Training Contracts (PRE2018-08383) was funded from the MCIN/AEI/10.13039/501100011033 and by the European Social Fund "ESF Investing in your future".

Acknowledgements

I would like to express my deepest gratitude to all the people who have contributed to the exciting and joyous journey that this doctoral thesis has represented.

First of all, I would like to start with my Ph.D. supervisor, Dr. María del Mar del Pino Yanes, whom I thank for the opportunity to conduct this work. Thank you profoundly for your dedication and time, as well as the passion you have introduced me with to Genetic Epidemiology.

Moreover, I would like to convey my gratitude to the members of the Genetics Area from the Department of Biochemistry, Microbiology, Cell Biology, and Genetics at the University of La Laguna (ULL) for your willingness to share your time and knowledge. Particularly, I would like to thank all my current and past colleagues from the Genomics and Health group, for your inspiration and assistance during this learning curve.

In addition, I would like to dedicate my sincere gratitude to all the collaborators involved in the completion of this doctoral thesis for your effort and dedication. In special, I would like to thank the members of the Asthma Collaboratory at the University of California San Francisco (UCSF) for their patience while answering my questions. Likewise, I would like to thank Dr. Erik Melén for hosting me in his research group at the Karolinska Institutet (KI) for the International stay, which I consider to be a tremendously enriching experience.

Finally, I will be eternally grateful to my wonderful family, blood or not, for the invaluable opportunity to learn with you every day. Especially, to my brave companions that keep showing me the wonders of *Towanda*.

Thank you all.

Abstract

Asthma exacerbations are episodes of symptoms worsening (i.e., dysnea, cough, wheezing or chest tightness) that require urgent intervention to prevent a serious outcome. These usually involve unexpected asthma care, hospitalizations and/or oral corticosteroids use, and are influenced by a complex interplay of genetic, environmental, and behavioural factors. Due to the clinical and socio-economic burden of asthma exacerbations, there is a critical need to establish potential biomarkers to guide prevention. This doctoral thesis aimed to identify genetic variants involved in asthma exacerbations through multiple genomic approaches. A literature search for previous genetic associations for asthma exacerbations was conducted in order to detect relevant loci to validate in independent populations with different ancestral backgrounds. Moreover, in order to explore genetic variation on six genomic regions harbouring genes whose combined sputum gene expression had been previously shown to predict asthma exacerbations, a candidate-gene association study of asthma exacerbations was conducted. Additionally, a genome-wide association study of asthma with severe exacerbations in ethnically diverse children and youth and a multi-ancestry meta-analysis of genome-wide association studies of asthma exacerbations revealed three novel genetic loci implicated in asthma exacerbations. Likewise, the exploration of genomic variation by leveraging local ancestry in a context of admixed populations with differential asthma exacerbations rates uncovered two genetic loci for severe asthma exacerbations that may exert ethnic-specific effects. Finally, the functional effects of the genetic variation on DNA methylation and gene expression were investigated using public online data bases and whole-blood data from ethnically diverse populations. Our findings provided insights into the pathogenesis of asthma exacerbations and prioritised potential targets of research moving forward precision medicine, including the *DNASE1L3*, *LINC01913*, *PKDCC*, *EXTL2*, *PANK1*, *DPYSL3*, and *SCGB3A2* genes.

Table of contents

Nomenclature **xv**

1 Introduction **1**

1.1 Asthma 1

1.1.1 Definition 1

1.1.2 Epidemiology 1

1.1.3 Diagnosis 2

1.1.4 Pathology and pathophysiology 3

1.2 Asthma exacerbations 5

1.2.1 Definition 5

1.2.2 Socio-economic burden 8

1.2.3 Risk factors 8

1.2.4 Triggers 11

1.2.5 Management 12

1.2.6 Genetics 15

1.2.6.1 Candidate-gene studies 18

1.2.6.2 Genome-wide association studies 19

1.2.6.3 Pharmacogenomic studies 20

1.2.6.4 Next-generation sequencing-based approaches 21
1.2.6.5 Admixture mapping 21

2 Hypotheses and aims 25

3 Materials and methods 27

3.1 Literature mining of genetic loci for asthma exacerbations 27
3.2 Study populations 27
3.3 Trait definition 31
3.4 Generation of -omic data, quality control and processing 32
3.4.1 Genetic variation data 32
3.4.2 DNA methylation data 34
3.4.3 Gene expression data 34
3.5 Statistical and bioinformatic analyses 35
3.5.1 Admixture mapping analyses 35
3.5.2 Allelic testing, meta-analysis, and independent variant assessment 35
3.5.3 Genetic associations shared with other traits 36
3.5.4 Assessment of functional effects 36

4 Chapters 39

4.1 Role of genomics in asthma exacerbations 39
4.2 A Deoxyribonuclease 1 Like 3 genetic variant associates with asthma exacerbations 46
4.3 Genome-wide association study reveals a novel locus for asthma with severe exacerbations in diverse populations 60
4.4 Admixture mapping of severe asthma exacerbations loci in children and youth 72

4.5 Multi-ancestry genome-wide association study of asthma exacerbations . 102

5 General discussion 137

5.1 Insights into the genetics and genomics of asthma exacerbations 137

5.2 Prioritization of *DNASE1L3* in asthma exacerbations 139

5.3 *LINC01913* and *PKDCC* as potential targets in asthma exacerbations . . 141

5.4 Leveraging local ancestry to detect genetic signals 144

5.5 *VCAM1*/*EXTL2* and *PANK1*: novel loci for asthma exacerbation 147

5.6 Study limitations 150

5.7 Future directions in the genomics of asthma exacerbations 152

6 Conclusions 155

References 157

Nomenclature

Abbreviations

1KGP 1000 Genomes Project

95%CI 95% confidence interval

AE Asthma exacerbation

ATS American Thoracic Society

AUC Area under the curve

BTS British Thoracic Society

CAAPA Consortium on Asthma among African-ancestry Populations in the Americas

CEU Utah residents with Northern and Western European ancestry

CHB Han Chinese in Beijing, China

CHD Chinese in Metropolitan Denver, Colorado

CoA Coenzyme A

COPD Chronic obstructive pulmonary disease

COVID-19 Coronavirus disease 2019

DNA Deoxyribonucleic acid

EAACI European Academy of Allergy and Clinical Immunology

eQTL Expression quantitative trait locus

eQTM Expression quantitative trait methylation

ER Acute asthma care, including Emergency room visits

ERP29 Endoplasmic reticulum protein 29

ERS European Respiratory Society

FEV_1 Forced expiratory volume in one second

FVC Forced vital capacity

GAGs Glycosaminoglycans

GTEx Genotype-Tissue Expression

GWAS Genome-wide association study

HRC Haplotype Reference Consortium

ICS Inhaled corticosteroids

ICU Intensive care units

IFN Interferon

IgE Immunoglobulin E

IL Interleukin

LABA Long-acting beta2-adrenergic receptor agonist

LD Linkage disequilibrium

LPS Lipopolysaccharide

LTRA Leukotriene receptor antagonist

MAF Minor allele frequency

MARCO Macrophage scavenger receptor with collagenous structure

meQTL Methylation quantitative trait locus

NAEPP National Asthma Education and Prevention Program

NAM Native American

NETs Neutrophil Extracellular Traps

NGS Next-generation sequencing

NOX2 Nicotinamide adenine dinucleotide phosphate oxidase 2

OCS Oral corticosteroids

OR Odds ratio

PEER Probabilistic estimation of expression residuals

PEF Peak expiratory flow

PRS Polygenic risk scores

QTL Quantitative trait locus

RAC1 Ras-related C3 botulinum toxin substrate 1

RNA Ribonucleic acid

ROC Receiver operating characteristic

RV-C Rhinovirus C

SABA Short-acting beta2-adrenergic receptor agonist

SIGN Scottish Intercollegiate Guidelines Network

siRNA Small interfering ribonucleic acid

SMAD Mothers against decapentaplegic homolog

SNP Single nucleotide polymorphism

STAT1 Signal transducer and activator of transcription 1

T2 Type 2

Th2 T helper 2

TMM Trimmed mean of M values

UK United Kingdom

US United States

VEGFR Vascular endothelial growth factor

WGS Whole-genome sequencing

YRI Yoruba in Ibadan, Nigeria

Gene abbreviations

ADAM33 A disintegrin and metalloprotease Domain 33

ADRB2 Adrenoceptor beta 2

ALPL Alkaline phosphatase

APOBEC3B Apolipoprotein B mRNA Editing Enzyme Catalytic Subunit 3B

APOBEC3C Apolipoprotein B mRNA Editing Enzyme Catalytic Subunit 3C

C11orf30 EMSY transcriptional repressor, BRCA2 interacting

CACNA2D3 Calcium voltage-gated channel auxiliary subunit alpha2delta 3

CDHR3 Cadherin related family member 3

CEACAM3 Carcinoembryonic antigen-related cell adhesion molecule 3

CLC Charcot-Leyden crystal galectin

CMTR1 Cap methyltransferase 1

CPA3 Carboxypeptidase A3

CRMP4 Collapsin response mediator protein 4

CRTAM Cytotoxic and regulatory T cell molecule

CTNNA3 Catenin alpha 3

CXCL8 C-X-C motif chemokine ligand 8

CXCR2 C-X-C motif rhemokine receptor 2

DNASE1L3 Deoxyribonuclease 1 like 3

DPYSL3 Dihydropyrimidinase like 3

EPHA7 Ephrin type-A receptor 7

EXTL2 Exostosin like glycosyltransferase 2

FCER2 Fc epsilon receptor II

FLJ22447 Uncharacterized LOC400221

FLNB Filamin B

GSDMB Gasdermin B

HIVEP2 Human immunodeficiency virus type I enhancer binding protein 2

HLA-DQB1 Major histocompatibility complex, class II, DQ beta 1

IKZF3 IKAROS family zinc finger 3

IL1B Interleukin 1 beta

IL1RL1 Interleukin 1 receptor like 1

IL4RA Interleukin-4 receptor alpha chain

KCNJ2-AS1 Potassium Inwardly Rectifying Channel Subfamily J Member 2 Antisense RNA 1

LINC01913 Long intergenic non-protein coding RNA 1913

LTBP1 Latent transforming growth factor beta binding protein 1

LY96 Lymphocyte Antigen 96

MUC5AC Mucin 5AC

PANK1 Pantothenate kinase 1

PTGER4 Prostaglandin E receptor 4

RAD50 RAD50 homolog, double strand break repair protein

SCGB3A2 Secretoglobin family 3A member 2

SGK493 Protein kinase-like protein SgK493

TBX3 T-Box transcription factor 3

TGFB1 Transforming growth factor beta 1

TSLP Thymic stromal lymphopoietin

VCAM1 Vascular cell adhesion molecule 1

VDR Vitamin D receptor

VLK Vertebrate lonesome kinase

WDR36 WD repeat domain 36

WNT5A Wnt family member 5A

Study population acronyms

ALLIANCE ALL Age Asthma Cohort

ALSPAC Avon Longitudinal Study of Parents and Children

BAMSE Children Allergy Milieu Stockholm an Epidemiological Study

followMAGICS Follow-up phase of the Multicenter Asthma Genetics in Childhood Study

GALA II Genes-environments and Admixture in Latino Asthmatics

GEMA Genomics and Metagenomics of Asthma Severity

GoSHARE Genetic of Scottish Health Research Register

HPR Hartford-Puerto Rico

INMA INfancia y Medio Ambiente

MEGA Mechanism underlying the genesis and evolution of asthma

MGI Michigan Genomics Initiative

PACMAN Pharmacogenetics of Asthma Medication in Children: Medication with Anti-inflammatory effects

PAGES Paediatric Asthma Gene-Environment Study

PASS Pharmacogenetics of Adrenal Suppression study

SAGE Study of African Americans, Asthma, Genes, and Environments

SCAALA Social Changes, Asthma and Allergy in Latin America

SCSGES Singapore Cross Sectional Genetic Epidemiology Study

U-BIOPRED Unbiased Biomarkers for the Prediction of Respiratory Disease Outcomes

UK Biobank United Kingdom Biobank

1

Introduction

1.1 Asthma

1.1.1 Definition

Asthma is a common heterogeneous disease of the respiratory system characterized by chronic inflammation and variable airflow obstruction that may become reversible or persistent with the progression of the disease. Asthma symptoms may manifest by cough, wheeze, breathlessness, and chest tightness (Global Initiative for Asthma, 2021a).

1.1.2 Epidemiology

The common risk factors for asthma development are age, gender, genetic ancestry and composition, allergic comorbidities, socioeconomic status, lifestyle factors (e.g., sedentary behaviour and obesity) and environmental factors, including early-life exposures. Some of the environmental factors comprise viral infections, tobacco, air pollution, or allergen exposure (Global Initiative for Asthma, 2021a; British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2019; Beasley et al., 2015).

Asthma affects 350 million people worldwide (GBD 2015 Chronic Respiratory Disease Collaborators, 2017), and it is estimated to impact up to 400 million people by 2025 (Masoli et al., 2004). It causes 250,000 deaths yearly (D'Amato et al., 2016) and generates a large burden on quality of life, physical and social functioning, work productivity and/or school attendance and increases the health care system expenditures (Nunes et al., 2017; Gruffydd-Jones et al., 2019; Oren and Martinez, 2020; Global

Initiative for Asthma, 2021a; Song et al., 2021). In fact, direct costs driven by asthma therapies and healthcare utilization as well as indirect costs due to days of school/work absenteeism are associated with asthma severity (Nunes et al., 2017; Song et al., 2020).

Patterns of asthma prevalence and incidence differ across geographical locations, with asthma prevalence ranging from 0.19% in China to 21% in Australia (To et al., 2012; Sears, 2014), and by ethnic groups. In particular, admixed populations with African ancestry, such as Latinos/Hispanics and African Americans, show higher rates of asthma susceptibility, morbidity, and mortality than non-Hispanic whites (Akinbami et al., 2012, 2014; Centers for Disease Control and Prevention, 2018). Interestingly, differences in asthma susceptibility are also observable among ethnic subgroups, with higher rates of asthma among Puerto Ricans compared to Mexicans, which could be related to differences in their genetic composition due to the specific historical processes of admixture (Hernandez-Pacheco et al., 2016; Bryc et al., 2015). Although asthma prevalence has increased since the 1990s, this trend appeared to have plateaued in many high-income countries in contrast to low- and middle-income countries (Dharmage et al., 2019), where most asthma-related deaths occur (Dharmage et al., 2019).

Although asthma is the most common chronic condition in children, it can develop at any stage of life. In particular, adults show rare asthma remission and higher rates of asthma mortality compared to children (Dharmage et al., 2019; Trivedi and Denton, 2019). Asthma prevalence and incidence also vary by biological sex in an age-dependent manner. During childhood, females exhibit lower asthma prevalence and incidence, but this trend reverses at adolescence (Shah and Newcomb, 2018).

1.1.3 Diagnosis

The clinical manifestations that characterize asthma are also observable in other chronic respiratory diseases; therefore, a detailed clinical assessment is required to establish a differential diagnosis of asthma. The presence of asthma symptoms and variable airflow obstruction are determined based on questionnaires, physical examination, pulmonary function tests, and measurements of other biomarkers (i.e., blood eosinophils count or fractional exhaled nitric oxide levels for eosinophilic inflammation) (British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2019; Global Initiative for Asthma, 2021a).

Pulmonary function tests comprise several measures, including forced expiratory flow in one second (FEV_1), forced vital capacity (FVC), their ratio (FEV_1/FVC), peak expiratory flow (PEF) (British Thoracic Society and Scottish Intercollegiate Guidelines Net-

work, 2019; Global Initiative for Asthma, 2021a). Since airflow obstruction in asthma is variable, normal pulmonary function tests values in a period of asymptomaticity should not exclude the diagnosis of asthma. This reversibility can be addressed by a bronchodilator response test, where lung function is measured before and after the administration of a short-acting bronchodilator (British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2019; Global Initiative for Asthma, 2021a). However, spirometry might not offer reliable results due to the challenges of performing the technique in children younger than six years old (Beydon et al., 2007). In this context, recurrent episodes of wheezing, cough, dyspnea or family history could contribute to the diagnosis (British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2019; Global Initiative for Asthma, 2021a). The presence of some of these symptoms may indicate a respiratory infection. Therefore, it is relevant to assess the duration of the symptoms under evidence of confirmed respiratory infection, the number of annual episodes and the presence of symptoms outside infection episodes (Global Initiative for Asthma, 2021a). In addition, complementary techniques that assess the presence of eosinophilic inflammation or atopy along with a therapeutic trial with asthma medication to assess the control of symptoms could contribute to exclude or confirm the diagnosis of asthma (British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2019; Global Initiative for Asthma, 2021a).

1.1.4 Pathology and pathophysiology

The diagnosis of asthma acts as an umbrella that comprises a plethora of clinical manifestations. Some of the most frequent and important asthma phenotypes are allergic asthma, non-allergic asthma, adult-onset asthma or obesity-related asthma (Global Initiative for Asthma, 2021a). Investigating specific asthma phenotypes and/or their related traits, such as Immunoglobulin E (IgE) levels or treatment response, in homogeneous subgroups of patients may lead to the identification of better biomarkers with clinical utility.

The complex heterogeneity of asthma may be explained by different symptoms underlying distinct pathophysiological mechanisms. Most patients can be stratified based on the population of T helper 2 (Th2) cells, type 2 (T2) cytokines, such as interleukins (IL) (e.g., IL-4, IL-5, IL-9 and IL-13) and eosinophilia. These mediators induce IgE production in basophils and mast cells, which plays a relevant role in the response against allergens, and the subsequent activation and accumulation of eosinophils in the sputum or peripheral blood. In turn, the eosinophils release effectors that promote diverse processes that culminate in airway inflammation and remodeling. The 'T2 high' subtype is

characterized by T2 innate lymphoid cells that release IL-5 and IL-3 (Yokoyama, 2019). Consecutively, IL-5 promotes recruitment and activation of eosinophils (Rothenberg and Hogan, 2006), whereas IL-13 induces mucus hypersecretion and airway inflammation (Hershey, 2003). The success of most asthma therapies lies in targeting the 'T2 high' specific profile, which is present in 50-80% of patients with asthma (Yokoyama, 2019). The definition of the 'T2 low' or non-T2 subtype is not clearly established, therefore, it is usually identified by the absence of a T2 high signature. The T2 low profile may be accompanied by neutrophilic or paucigranulocytic inflammation. Moreover, two major immune pathways can be observed in non-T2 asthma. The first one involves type 3 immunity mediated by IL-17, and the second one type 3 immunity mediated by interferon (IFN). Patients with non-T2 asthma may also exhibit systemic inflammation with a high IL-6 signature (Fitzpatrick et al., 2020; Sze et al., 2020).

Several physiological events underlie asthma symptoms. Bronchoconstriction is one of the major pathophysiological events in asthma, and can be triggered by different stimuli such as allergens, irritants, infections, cold air, or exercise. Tightening of the airway smooth muscles leads to airway narrowing and the subsequent airflow limitation (U.S. National Heart, Lung, and Blood Institute of the National Institutes of Health, 2007). Notably, bronchoconstriction can promote airway remodeling independently of inflammation (Grainge et al., 2011). These changes in the smooth muscle may enhance airway sensitivity and reactivity to direct or indirect stimuli. Airway hyperresponsiveness is frequently present under the context of chronic inflammation, although it can be triggered by acute transient inflammation (Meurs et al., 2008) and independently of an inflammatory allergic response (Kalotas et al., 2021).

The respiratory epithelium is not only the frontline physical barrier but a functional and immunologic boundary to protect the host against allergens, pollutants or microbes. Briefly, the airway epithelium recognizes patterns associated with these damaging agents and triggers the secretion of pro-inflammatory mediators and the recruitment of immunoglobulins. Moreover, patients with asthma may exhibit airway wall thickening, mucus hypersecretion and hyperplasia of mucous glands in the airway epithelium, which leads the deregulation of the machinery of mucociliary clearance (Frey et al., 2020; Niimi et al., 2000). Airway remodeling comprises a compendium of structural changes in the airways, including thickening of the reticular basement membrane, subepithelial fibrosis, airway smooth muscle hypertrophy and hyperplasia, enhanced airway vascularity, mucus hypersecretion, and mucous gland hyperplasia (U.S. National Heart, Lung, and Blood Institute of the National Institutes of Health, 2007). Airway epithelial cells regulate the morphological status of airway smooth muscle via cytokines and contractile agonists. Moreover, they also chemically induce fibroblasts

and myofibroblasts to prompt the secretion of extracellular matrix proteins that contribute to structural changes in the airways and fibroblast-to-myofibroblast transition, eventually leading to subepithelial fibrosis (Hough et al., 2020). It is hypothesized that airway remodeling alters the threshold above which asthma flare-ups are triggered through narrowing of the airways, which also affects lung function (Krings et al., 2021; Winkler and Frey, 2021). Although airway remodeling can occur as a repair response to chronic inflammation, it can also occur due to abnormal lung development independently of inflammation processes (Grainge et al., 2011; Fehrenbach et al., 2017).

Airway epithelial barrier deregulation enhances the translocation of exogenous damaging agents, usually accompanied by chronic local inflammation. Interestingly, microbial dysbiosis may alter relevant functions in the immune response. For example, the gut decline in taxa that produces short-chain fatty acids increases susceptibility to asthma (Boutin et al., 2021) and can lead to alterations in the inflammatory and immune response via the gut-lung axis (Enaud et al., 2020). Despite the fact that most of the microbes present in the lung are bacteria, these interplay with other microbial communities to modulate their roles (Liu et al., 2020; Boutin et al., 2021).

1.2 Asthma exacerbations

1.2.1 Definition

Asthma patients can experience the worsening of their usual day-to-day asthma symptoms (coughing, wheezing, dyspnea, chest tightness, and/or nighttime awakenings) that can lead to absences to school or work, seeking acute asthma care, use of oral corticosteroids (OCS), hospitalizations, or even death (Global Initiative for Asthma, 2021a). These events, known as asthma exacerbations (AEs) or flare-ups, have been heterogeneously defined in the literature.

In order to standardize the definition of AEs, a task force designated by the American Thoracic Society and the European Respiratory Society (ATS/ERS) convened on the definition of AEs based on their severity. Severe AEs were defined by asthma-related hospitalizations or emergency room visits requiring corticosteroids or, alternatively, the use of corticosteroids for at least three days to prevent a serious outcome. Moderate exacerbations were defined by worsening of the usual day-to-day symptoms or lung function and/or increased rescue bronchodilator use for at least two days. These events usually require a change in treatment but not the use of oral corticosteroids, hospitalizations, or emergency room visits (Reddel et al., 2009). More recently, a task force

organized by the ERS and the European Academy of Allergy and Clinical Immunology (EAACI) suggested that the use of systemic corticosteroids for at least five days might be more clinically relevant in severe AEs (Bourdin et al., 2019). However, this is a subjective utilization-based criteria that relies on the perception of the patient and/or the clinicians rather than in clinical or physiological parameters.

Given the complexity of defining AEs, some guidelines aim to establish more objective criteria to determine the severity of the exacerbations that could guide pharmacological management of the exacerbation (**Table 1**). The Global Initiative for Asthma (GINA) defines exacerbations as "episodes characterised by a progressive increase in symptoms of shortness of breath, cough, wheezing or chest tightness and progressive decrease in lung function, i.e., they represent a change from the patient's usual status that is sufficient to require a change in treatment" (Global Initiative for Asthma, 2021a). This guideline provides separate definitions of AEs for children aged five years or younger and for children aged 6 years or more and adults, based on symptoms, clinical findings, and lung function. In the former age group, lung function is not considered because of the difficulty of applying this technique reliably (Beydon et al., 2007). Moreover, the British Thoracic Society and the Scottish Intercollegiate Guidelines Network (BTS/SIGN) also clarifies the principles to determine the severity of the exacerbation based on symptoms and clinical findings.

Table 1. Comparison of the BTS/SIGN 2019 and GINA 2021 criteria for the definition of AEs per severity and age group.

	BTS/SIGN		GINA	
Severity	**Children**	**Adults**	**Children (≤5 years)**	**Children ≥6 years and adults**
Moderate	· Able to talk in sentences · SpO_2 ≥92% · PEF ≥50% best or predicted · HR (bpm) ≤140* or ≤125† · RR ≤40/min* or ≤30/min†	· Increasing symptoms · PEF >50–75% best or predicted · No 'severe' symptoms	· Unaltered consciousness · SaO_2 > 95% · 'Talks in sentences'‡ · HR <100bpm · RR ≤40 breaths/minute · No central cyanosis · Variable wheezing intensity	· 'Talks in phrases','Prefers sitting to lying', 'Not agitated' · Accessory muscles not used · RR increased but <30/min · HR ranging 100–120bpm · SaO_2 in air 90–95% · PEF >50% predicted or best
Severe	· Can't complete sentences in one breath or too breathless to talk or feed · SpO_2 <92% · PEF 33–50% best or predicted · HR (bpm) >140* or >125† · RR >40/min* or >30/min†	Any one of: · PEF 33–50% best or predicted · RR ≥25/min · HR ≥110bpm · Inability to complete sentences in one breath	· Being agitated, confused or drownsy · SaO_2 < 92% · 'Talks in words'‡ · HR (bpm) >180 (0-3y) or >150 (4-5y) · RR >40 per minute · Central cyanosis · Silent chest under auscultation	· 'Talks in words', 'sits hunched forward', 'agitated" · RR >30/min · HR >120bpm · SaO_2 <90% on air · Use of accessory muscles · PEF ≤50% predicted or best
Life threatening		Severe asthma and any of: · Altered conscious level · Exhaustion · Arrhythmia · Central cyanosis · Hypotension · Poor respiratory effort · Silent chest under auscultation · PEF <33% best or predicted · SpO_2 <92% · PaO_2 <8 kPa · 'Normal' $PaCO_2$ (4.6-6.0 kPa)		
Near fatal		· Raised $PaCO_2$ and/or requiring mechanical ventilation		

*Children aged 0-5 years old; †Children over 5 years; ‡Taking into account the developmental stage. Abbreviations: bpm: Beats per minute; HR: Heart rate; kPa: kilopascals; PaO_2: partial arterial pressure of oxygen; $PaCO_2$: partial arterial pressure of carbon dioxide; PEF: Peak expiratory flow; RR: Respiratory rate; SpO_2: oxygen saturation measured by a pulse oximeter.

1.2.2 Socio-economic burden

Similar to asthma prevalence, exacerbation rates vary widely among geographical locations and ethnic groups. In the United States, African Americans and Puerto Ricans show an increased number of hospitalizations, emergency department, and/or urgent care center visits due to asthma compared to Whites (Akinbami et al., 2012; Oraka et al., 2013; Akinbami et al., 2014; Rosser et al., 2014). In Europe, the rates of AEs are higher in Southern European countries (Henneberger et al., 2010; Engelkes et al., 2020a). Overall, the annual rate of AEs in the United States and Europe is around 0.1-0.2 per patient with asthma (Suruki et al., 2017; Bloom et al., 2019; Engelkes et al., 2020a; de Roos et al., 2021), increases with the severity of the disease (Suruki et al., 2017; Engelkes et al., 2020a), and reaches 0,67 adults treated with a high dose of inhaled corticosteroids (Schatz et al., 2014).

AEs increase the risk of mortality among asthma patients (Ali et al., 2013; Engelkes et al., 2020b). The in-hospital mortality following admission due to AEs is approximately 0.2-1.5% (Krishnan et al., 2006; Strid et al., 2013; Engelkes et al., 2020b). Since there are no significant differences in the in-hospital mortality rate among African Americans and European-descent patients with asthma, the elevated overall mortality among African Americans may be explained due to out-of-hospital deaths (Krishnan et al., 2006).

The burden of AEs includes reduced quality of life in patients with moderate, severe or difficult-to-treat asthma (Lloyd et al., 2007; Luskin et al., 2014; Chipps et al., 2018), related to the severity of the exacerbation (Chipps et al., 2018), and also extends to alterations of the psychological state of the caregivers (Rastogi et al., 2013). Nowadays, management of AEs causes high average annual expenditures for total and asthma-related healthcare (Ivanova et al., 2012) that increases with the severity of the exacerbation (Suruki et al., 2017), particularly in secondary care (Lane et al., 2006).

1.2.3 Risk factors

The complexity of AEs could be addressed from a perspective in which different elements interplay. Risk factors predispose to a greater susceptibility to these events, whereas exposure to certain agents or stimuli known as "triggers" may precipitate exacerbations. Additionally, the presence of comorbidities can affect the progression or severity of the episode. In fact, some factors can belong to different categories (e.g., smoking can trigger and predispose to AEs) (Ioachimescu and Desai, 2019).

Numerous risk factors for AEs have been described, but not all of them are strong predictors for AEs or have been validated across cohorts. Risk factors for AEs have been discussed in depth elsewhere (Rosser et al., 2014; Puranik et al., 2017; Martin et al., 2020; Navanandan et al., 2021). The fact that the strongest risk factor and best predictor of exacerbations to date is having had a previous one, particularly in the last year, suggests that susceptibility to AEs is determined by heterogeneous mechanisms where genetics, behavioural, and environmental factors interplay (Ramsahai et al., 2019; Martin et al., 2020; di Palmo et al., 2021). The genetics of AEs will be discussed in epigraph 1.2.6.

AEs occur more frequently in patients with severe asthma, albeit patients with mild asthma, which account up to 70% of asthmatics, also suffer exacerbations (Nunes et al., 2017; Suruki et al., 2017; Bloom et al., 2019; Global Initiative for Asthma, 2021a). Despite asthma severity and control are recognized as a risk factors for AEs, their predictive capability is limited (Martin et al., 2020; Navanandan et al., 2021). In fact, predictor of persistent asthma symptoms differs from those for AEs per se (Wu et al., 2011).

In the United States, minority groups at high risk of asthma exhibit differential exacerbation rates (Akinbami et al., 2012; Oraka et al., 2013; Akinbami et al., 2014; Rosser et al., 2014). Additionally, African American pediatric patients with AEs are more likely to incur on intensive care units (ICU) use and longer lengths of stay compared to European-descent patients in the context of public insurance (Silber et al., 2017; Grunwell et al., 2018). Similarly, ICU admissions are higher among African Americans than in Latinos (Lee et al., 2020a). Interestingly, African ancestry has been associated with higher risk of AEs defined as burst oral steroid use, hospitalization, and/or emergency department visits in the previous year among African American children and young adults (Rumpel et al., 2012) and in African-descent individuals in the United States (Grossman et al., 2019). However, an 18-month longitudinal study did not find significant association between African ancestry and the risk of AEs in the past three months prior to each follow-up visit in African American children and adults (Flores et al., 2012). Moreover, African ancestry was not significantly associated with acute asthma care and/or hospitalization in Puerto Rican children with asthma (Brehm et al., 2012). Although lower socioeconomic status is associated with increased risk of AEs, it does not fully account for the association between genetic ancestry and asthma outcomes (Grossman et al., 2019; Ramsahai et al., 2019).

Younger age is also associated with higher risk for severe AEs in pediatric asthma (Engelkes et al., 2020a), whereas older age among adults is likely to increase the risk for AEs (Kang et al., 2018; Racine et al., 2021). Nonetheless, pediatric patients with asthma are more likely to suffer exacerbations than adults (Akinbami et al., 2012).

The risk of exacerbations is heightened by comorbidities such as chronic rhinitis and gastroesophageal reflux disease (Martin et al., 2020). Kang et al. (2018) found that age and comorbidities are relevant risk factors for AEs in mild asthma, whereas drug adherence and prior hospitalizations are most important in severe asthma. However, rhinosinusitis and gastroesophageal reflux disease were associated with the frequency of exacerbation in an asthma cohort where 60% of the patients had severe asthma (Denlinger et al., 2017). Some studies have also found that obese individuals are affected by an increased risk for AEs (Black et al., 2013; Schatz et al., 2013; Denlinger et al., 2017; Longo et al., 2018), and the response to budesonide measured as exacerbations (Forno et al., 2011).

A similar age-dependent trend for the effect of sex on asthma is observed for AEs. Adult females exhibit higher odds of experiencing AEs compared to their male counterpart while boys are more likely to undergo exacerbations than young girls (Tattersfield et al., 1999; Sears, 2008; Kynyk et al., 2011; Patel et al., 2014). In fact, adult women who present at the emergency department with acute asthma are more likely to be admitted than males despite having better lung function (Trawick et al., 2001; Chasm et al., 2015; Kang et al., 2018). Regardless of the age group, females are more susceptible to have longer hospital stays (Hasegawa et al., 2015).

The presence of T2 airway inflammation, measured as high blood eosinophil count (Malinovschi et al., 2013; Price et al., 2015; Denlinger et al., 2017; di Palmo et al., 2021), is also recognized as a risk factor for AEs. In fact, sputum and blood eosinophil counts are predictors of life-threatening or moderate-to-severe AEs (Esmaeilzadeh et al., 2021). Semprini et al. (2018); Briggs et al. (2021) found that T2 biomarkers affected the risk of severe exacerbations in severe refractory asthma rather than in mild or moderate asthma. In regards to the effect of exhaled nitric oxide on AEs, heterogeneous evidence has been described for this biomarker for eosinophilic airway inflammation (di Palmo et al., 2021). Additionally, neutrophils play an important role at Th2 cytokines-mediated eosinophil recruitment in rhinovirus-induced AEs (Busse, 2017). Interestingly, neutrophilia has been associated with more frequent exacerbations, although the event's severity is lower compared to those reported in asthma patients with eosinophilia (Syabbalo, 2020). In addition, the neutrophil–lymphocyte ratio, an indicator of systemic inflammation is also associated with increased risk for AEs (Esmaeilzadeh et al., 2021)

Patterns of medium- and long-term lung function suggested that AEs have been associated with an accelerated deficit of lung function, measured as peak expiratory flow and pre and/or post-bronchodilator FEV_1, in adults and children (Bai et al., 2007; O'Byrne et al., 2009; O'Brian et al., 2012; Briggs et al., 2021; Major et al., 2021). A

recent longitudinal study that followed European patients with asthma from infancy through childhood found that exacerbators had increased airway resistance and exhibited a decline in maximal mid-expiratory flow compared to non-exacerbators. However, no significant differences were found for pre-bronchodilator FEV_1, and no significant changes on the lung function trajectories were found prior and after exacerbations. All this together suggests that airway obstruction is likely to mediate the risk of exacerbations, instead of being the consequence of these events (Hallas et al., 2020).

1.2.4 Triggers

Triggers for AEs can be differentiated based on their allergic or non-allergic nature (Ioachimescu and Desai, 2019). Allergic triggers include outdoor and indoor exposures. Regarding the outdoor allergen exposure, it is worth noting that the spring peaks of asthma-related emergency room visits and hospitalization correlates with events of high grass pollen counts and humidity (Erbas et al., 2018; Silver et al., 2018). These acute episodes of thunderstorm asthma are likely to increase T2 airway inflammation. Indoor exposures including allergic sensitization to dust mites or pests are also associated with higher odds for AEs. Non-allergic triggers include respiratory infections, tobacco smoke, emotional factors, treatment adherence and air pollution (Ioachimescu and Desai, 2019).

Viral infections are recognized as the main trigger of AEs. In the pediatric population, rhinoviruses are the most frequent viral trigger. In particular, rhinovirus C (RV-C) is commonly detected in children with AEs in the back-to-school period, whereas rhinovirus A is the most prevalent in adults. Nonetheless, the influenza virus is the most commonly detected virus among adults with exacerbations, particularly in winter (Papadopoulos et al., 2011; Ramsahai et al., 2019). Although AEs may be triggered by fungal (e.g., *Aspergillus*) or bacterial agents, the role of the microbiota in AEs remains scarcely explored (Fazlollahi et al., 2018; McCauley et al., 2019; Ramsahai et al., 2019; Zhou et al., 2019).

Smoking exposure, either by self-smoking or second-hand smoking, is associated with higher odds of AEs (Ioachimescu and Desai, 2019). In Latinos and African Americans asthma, second-hand smoking has been associated with poor asthma control and increased risk for AEs even in children with low plasma cotinine levels (Neophytou et al., 2018)

Poor adherence to asthma controllers is associated with higher odds of AEs in children and adults (Engelkes et al., 2015). In fact, controller non-adherence in asthma is

high (Engelkes et al., 2015), which could be explained by several factors such as the patient's perceptions, concerns about side effects and social stigma (De Simoni et al., 2017). Moreover, the psychological status of the patient may also trigger AEs. Stress, anxiety, or depressive symptoms have been consistently associated with increased risk and frequency of AEs in adolescents and adults (Ritz et al., 2016; de Roos et al., 2021).

Increased exposure to asthma triggers is associated with the frequency and severity of the exacerbation and lower quality of life (Luskin et al., 2014; Ritz et al., 2016; Chipps et al., 2018). Interestingly, AEs declined during the coronavirus disease 2019 (COVID-19) regardless of sex, age group, ethnicity, or geographical location (Lam et al., 2021; Salciccioli et al., 2021; Shah et al., 2021). Individuals that usually worked outside of their home and those without T2 inflammation showed the greatest decrease (Salciccioli et al., 2021), which may indicate the benefit of trigger avoidance. This decrease may be also partially explained by improved self-management of the symptoms or increased in-home management of the exacerbation to avoid in-person health care contact. However, Salciccioli et al. (2021) found no relevant differences between remote and in-person care. Moreover, the trend remained significant after controlling for air quality values, which questions the relevance of air pollution as an AEs trigger (Salciccioli et al., 2021).

1.2.5 Management

Pre-hospital management of AEs focuses on the prevention of these events via pharmacological and non-pharmacological interventions. The latter type of intervention includes avoidance of trigger exposures, preventive vaccination (e.g., influenza virus vaccine), regular exercise (Murray et al., 2021), as well as educational initiatives and clinical monitoring. The role of the caregiver is essential in the prevention and management of these events. For example, the caregiver perception of the child's asthma control, independent of the guidelines assessment, predicted the risk for future exacerbations in low-income minority children with persistent asthma (Kochis et al., 2019). Moreover, a high percentage of the caregivers may not provide adequate in-home management of the events (Yeung et al., 2018). Educational interventions, guided self-management, and other preventive care (e.g., well care center visits) contribute to significantly decreasing the risk of AEs and their frequency (Rastogi et al., 2013; Utidjian et al., 2017; Hodkinson et al., 2020; Kenyon et al., 2020; Lang et al., 2020; Nanishi et al., 2021).

Pharmacological treatment in asthma considers two types of medication according to the time they exert their effects: reliever or rescue medications, which offers a quick

relief of the symptoms, and controller medications that require routine use for long-term effects on asthma control. The backbone of the reliever therapies in children aged 6 years or older, adolescents and adults according to current guidelines is a combination of inhaled corticosteroids (ICS) and formoterol, long-acting beta2-adrenergic receptor agonists (LABAs), or alternatively, use short-acting beta2-adrenergic receptor agonists (SABAs), which shows reduced power to control AEs compared to the first therapy (British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2019; Global Initiative for Asthma, 2021a). ICS amelioriate AEs characterized by eosinophilia, but has not shown significant effects in non-eosinophilic exacerbations (Jayaram, 2006). This is likely due to the reduction of T2 inflammatory gene expression via the glucocorticoid receptor (Ramadan et al., 2019). Additionally, LABA is thought to reduce the exacerbation via synergistic effects with ICS on airway inflammation and remodeling (Kelly et al., 2010). Alternatively, short-term administration of OCS could be used for quick amelioration of asthma symptoms, albeit side-effects are expected (British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2019; Global Initiative for Asthma, 2021a). The most commonly prescribed controller medications in asthma are ICS, LABAs and leukotriene receptor antagonists (LTRAs), although LABAs are only recommended in children aged six years or older (British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2019; Global Initiative for Asthma, 2021a).

In severe asthma, other treatments can be considered on a patient-by-patient basis when the standard approach does not provide adequate control. Long-acting muscarinic antagonists may be used as an alternative treatment in eighteen years or older patients (Global Initiative for Asthma, 2021a). Moreover, in adults with severe persistent asthma, the macrolide antibiotic azithromycin reduces the risk of exacerbations likely due to anti-infective and anti-inflammatory effects (Gibson et al., 2019). Biological therapies with different monoclonal antibodies targeting components of T2 inflammation have been shown to effectively reduce the risk of AEs for different subtypes of severe asthma (British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2019; Global Initiative for Asthma, 2021a). The most common monoclonal antibodies are omalizumab, mepolizumab, and dupilumab. Omalizumab joins to free IgE and, therefore, blocks the interaction with the IgE receptor in basophils and mast cells, and reduces the subsequent allergen-induced cell activation. In addition, blocking IgE enhances IFN-α response to viral infections, which could exert a protective effect against viral-induced AEs (Cardet and Casale, 2019). Mepolizumab, a monoclonal antibody against IL-5, reduces AEs, likely by decreasing the IL5-mediated recruitment of eosinophils (Shrimanker et al., 2018). Dupilumab reduces AEs in patients with and without eosinophilia via blockage of the action of IL-4 and IL-13 (Wenzel et al., 2016).

Currently, there is no evidence that dupilumab and mepolizumab may play a role in virus-induced exacerbations.

The practices for pharmacological interventions are standardized by international or national guidelines for asthma management that describe step-wise treatment approaches to reduce and control asthma symptoms, and prevent AEs. After initial assessment of the symptoms control, lung function, risk of exacerbations, medication side-effects, the appropriate therapeutic step is determined and re-assessed up to three months later (or before, based on the severity of the symptoms) if symptoms control is not achieved (Global Initiative for Asthma, 2021a).

Despite the evidence-based step-wise approaches proposed by these guidelines are constantly updated (British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2019; Global Initiative for Asthma, 2021a), this doctoral thesis includes studies with data from patients with asthma recruited in different studies worldwide that have been treated following specific version of the guidelines based on their recruitment period. These include BTS/SIGN 2014 (British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2014) and GINA 2006-2014 (Global Initiative for Asthma, 2021b) guidelines along with the Expert panel report 3 for the Diagnosis and Management of Asthma 2007 from the United States National Asthma Education and Prevention Program (NAEPP) (U.S. National Heart, Lung, and Blood Institute of the National Institutes of Health, 2007). **Figure 1** depicts the step-wise pharmacological strategy recommended by the BTS/SIGN 2014, which was followed in most of the studies included in this doctoral thesis.

Management of AEs in care facilities is guided by the severity of the exacerbation and differs by age group and clinical setting. GINA addresses management of exacerbations in primary care and acute care facilities, while BTS/SIGN differentiates between general practice, emergency room and hospital. After the initial assessment to determine the severity of the exacerbation, an immediate treatment is established. Then, the patients should be monitored frequently to assess their progress via lung function, symptoms and clinical findings. If the exacerbation recedes, then discharge, treatment modification and subsequent follow-ups are considered. When the exacerbation status does not improve, treatment should be continued and transferred to a high-level care if needed. Treatment of acute asthma flare-ups consider primarily therapies that attenuates airway obstruction and inflammation. If evidence of respiratory infection exists, antibiotics could also be considered. Apart from the medication mentioned above and/or adjustment on the dose of each therapy, other alternative and/or complementary treatments could be contemplated on a patient-basis. These include magnesium to promote smooth muscle relaxation and subsequent bronchodilation,

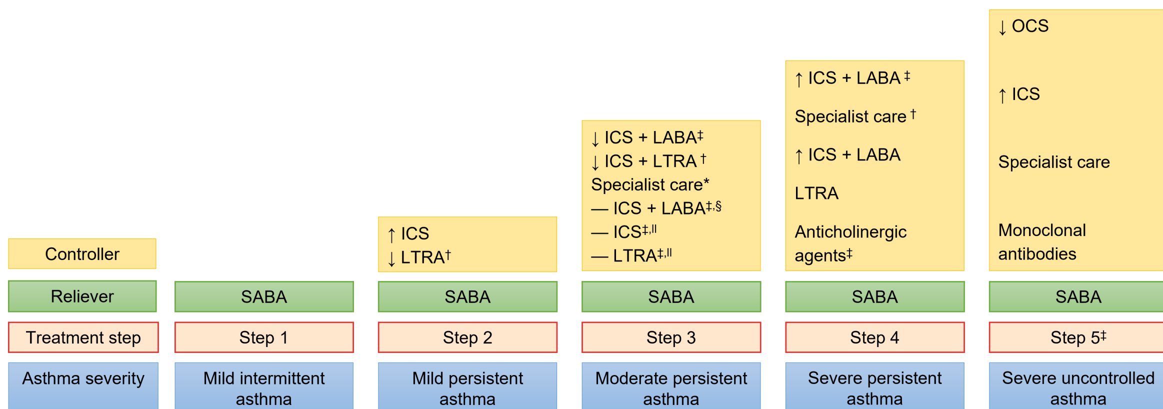


Figure 1. Step-wise pharmacological approach to control asthma symptoms and prevent exacerbations, according to the BTS/SIGN 2014 guidelines. Treatment dosages are represented by arrows as follows: high (↑), medium(—) and low (↓). * Children <2 years old; † Children <5 years old; ‡ Children ≥ 5 years old and adults; § Patients in LABA therapy who exhibit some benefits but inadequate control; ǁ Non-responders to LABA therapy. Based on British Thoracic Society and Scottish Intercollegiate Guidelines Network (2014). ICS: inhaled corticosteroids; LABA: long-acting beta2 agonists; LTRA: leukotriene receptor antagonists; SABA: short-acting beta2 agonists.

supplemental oxygen therapy, or epinephrine in case of anaphylaxis or angioedema (British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2019; Global Initiative for Asthma, 2021a).

1.2.6 Genetics

AEs constitute a huge burden on asthma patients, their caregivers and the health system (Nunes et al., 2017; Gruffydd-Jones et al., 2019; Oren and Martinez, 2020; Song et al., 2020; Global Initiative for Asthma, 2021a; Song et al., 2021). However, to date, the best predictor of AEs is having had one before, particularly in the previous year (Ramsahai et al., 2019; Martin et al., 2020; di Palmo et al., 2021). Therefore, the identification of subgroups of patients at high risk of AEs is a crucial and urgent need in the clinical setting. In this context, genetic research aims to understand the association of AEs and the genetic variability throughout the genome. Moving forward for therapeutics development and to the application of precision medicine into the clinical practice, identifying genetic markers associated with AEs that can be used to discriminate patients at risk of suffering these events for preventive purposes, as well as identify underlying biological mechanisms is a priority (di Palmo et al., 2021).

A large number of genetic loci for AEs has been uncovered by different genomic approaches (**Figure 2**). These comprised candidate-gene association studies, genome-wide association studies (GWAS) and combination with other -omics. Moreover, some of these focused on the role of gene-environment interactions on AEs or have analyzed AEs as an outcome of asthma treatment response.

Figure 2. Circos plot for genetic loci for AEs. The first track shows the heatmap for the minor allele frequency in the superpopulations from the KGP (from outer to inner: African admixed, American admixed, East Asian, European and South Asian populations). The second track shows the samples size per study type by color (Candidate-gene study: yellow; GWAS: red; Pharmacogenetic study: pink; Gene-environment interaction: green) and form (Candidate gene study: triangle; GWAS: circle; Combination with other omics: square).

Some of these studies aimed to untangle the regulatory effects of the relevant loci. Genetic variation may cause epigenetic changes, such as DNA methylation and histone modification, that alter DNA accessibility and chromatin structure. In turn, these epigenetic changes may result in alteration in gene and protein expression levels. Moreover, transcriptional and cotranscriptional processes controlled by genetic variation can also regulate gene and protein expression levels (**Figure 3**) (Pai et al., 2015; Ye et al., 2020).

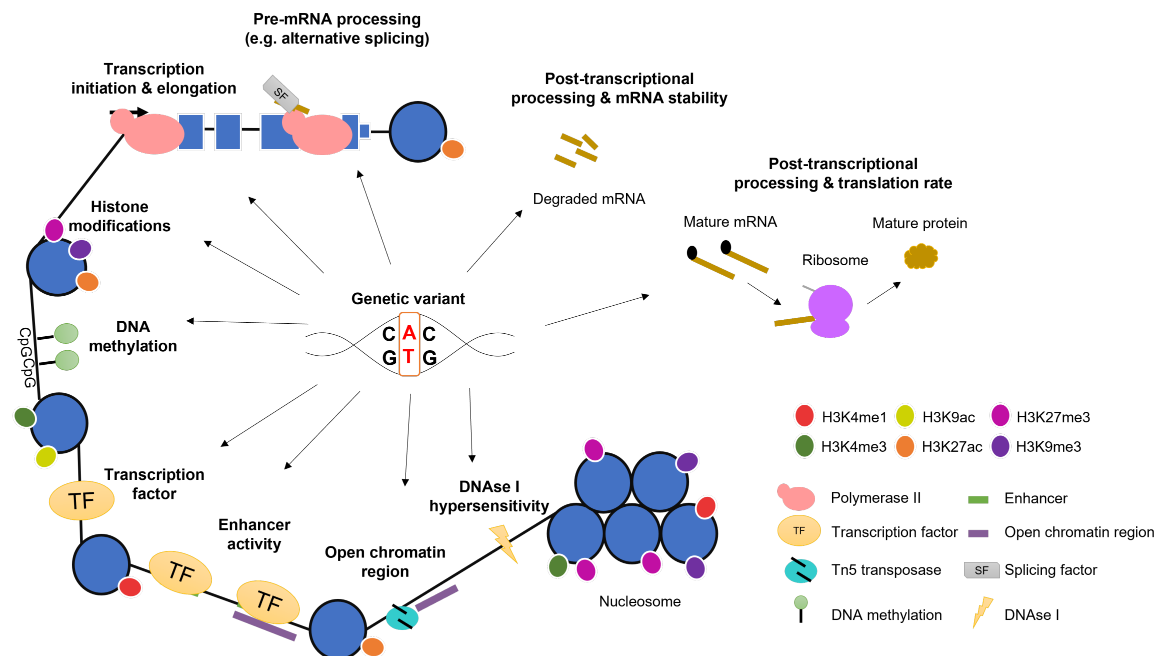


Figure 3. Effects of genetic variation on DNA methylation, gene and protein expression. First, genetic variants can exert epigenetic modifications via different mechanisms: i) Multiple types of histone modifications, such as (tri-) methylation of lysine 3, 9, or 27 of histone 3 (H3K4me, H3K4me3, H3K9me3, H3K27me3) and acetylation of lysine 9 and 27 of histone 3 (H3K9ac, H3K27ac); ii) DNA methylation modifications, which occurs more frequently at CpG sites; and iii) Regulatory elements (Transcription factors, enhancer activity, open chromatin regions, DNAse I hypersensitivity sites, ...). These changes may also modulate gene expression. Furthermore, genetic variants can regulate gene and protein expression following several mechanisms: iv) Transcriptional factor binding; v) Transcriptional elongation (by Polymerase III traveling rates); vi) pre-mRNA processing and modification; vii) Post-transcriptional mRNA processing, including mRNA degradation, effects in polyadenylation, and targeting by miRNAs; and viii) Post-transcriptional and translational mechanisms where mRNA expression levels may impact downstream protein products. Based on Pai et al. (2015); Ye et al. (2020). CpG: CpG site, a region of the genome where a cytosine nucleotide is followed by a guanine nucleotide in the linear sequence of bases along its 5' → 3' direction; mRNA: messenger RNA.

The role of genetic variation on DNA methylation and gene expression can be explored by means of quantitative trait locus (QTL) mapping. In particular, expression

quantitative trait loci (eQTL) analyses investigate the association of genotypes and variation in gene expression levels, whereas methylation quantitative trait loci (meQTL) analysis focus on the association of genotypes and DNA methylation at 5'-cytosine-phosphate-guanine-3' dinucleotide (CpG) sites. Besides, expression quantitative trait methylation (eQTM) analyses assess the association of DNA methylation and gene expression levels (**Figure 4**).

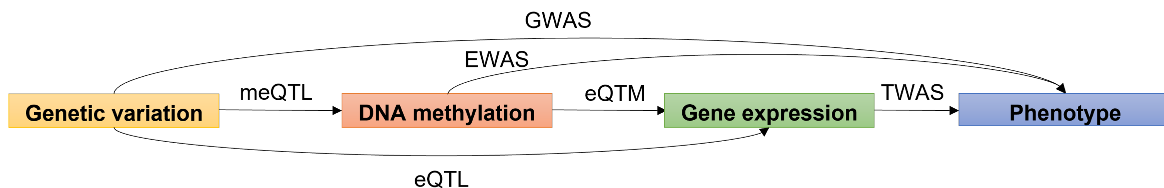


Figure 4. Combination of different -omic and clinical layers to understand the role of genetic variation on DNA methylation and gene expression. The association of genetic variation with a specific trait, DNA methylation or gene expression is assessed by GWAS, epigenome-wide association studies (EWAS) and transcriptome-wide association studies (TWAS). In addition, a regulatory genetic variants can exert effects over DNA methylation (meQTL) and/or gene expression (eQTL). Moreover, DNA methylation levels at a specific chromosomal position may regulate gene expression levels (eQTM).

1.2.6.1 Candidate-gene studies

Candidate-gene studies have been the most common strategy in the search of genetic predictors of AEs. These types of genetic analyses focus on a limited set of genetic variants located within or close to genes that are hypothesized to be involved in the trait of interest because of their previously known roles (Duncan et al., 2019). Although the catalog of human genetic variation is wide, candidate-gene studies have usually investigated single nucleotide polymorphisms (SNPs), which are changes of one nucleotide in the DNA sequence that occurs with a minor allele frequency (MAF) equal or higher than 1% in the general population (Ku et al., 2010).

Due to the hypothesis-driven nature of these studies, it is not surprising that most of candidate-gene studies of AEs have assessed genetic variation within genes previously implicated in childhood-onset asthma (e.g., the 17q12-21 locus) and/or asthma-related processes, such as type-2 inflammation (e.g., *IL-33*, IL4RA and *IL1RL1*), lymphocyte B growth and IgE production (*FCER2*), or treatment response (*ADRB2*), among others (di Palmo et al., 2021). However, this hypothesis-driven strategy is not exempted of limitations, including the identification of low effect variants, low likehood of revealing new biological mechanisms involved in the trait of interest, or lack of reproducibility (Duncan et al., 2019).

1.2.6.2 Genome-wide association studies

The advent of genotyping platforms that improved the genetic coverage throughout the genome led to the development of hypothesis-free GWAS. This analytical approach also relies on a procedure called 'imputation' to increase the genome-wide coverage based on haplotype reconstruction using the correlation between genetic variants (usually estimated from reference panels that gather the linkage disequilibrium [LD] structure from up to hundreds of thousands of individuals worldwide to date) (Quick et al., 2020).

The first genome-wide scan of AEs focused on the interaction of genetic variants with vitamin D levels on AEs in 403 non-Hispanic White subjects and trios (Du et al., 2012). The homozygous genotype for the A allele of rs2272094 (*CRTAM*) increased the rate of AEs in patients with low vitamin D levels (≤30 ng/mL) and *CRTAM* gene expression in the presence of vitamin D. The authors suggested that *CRTAM* may play a role through CD8+ and NKT cells during viral infections.

Additionally, a GWAS in Danish children younger than six years old compared 1,173 asthma patients with recurrent asthma-related hospitalizations and 2,522 non-asthma controls. This specific strategy comparing extreme phenotypes flagged rs6967330 (*CDHR3*) as a novel susceptibility locus for asthma with exacerbations (Bønnelykke et al., 2014). *CDHR3* is the only known receptor for RV-C and a driver of cilia differentiation in airway epithelium (Lutter and Ravanetti, 2019). The risk allele (A) for asthma with recurrent hospitalizations enhances *CDHR3* gene and protein expression, which in turn increases RV-C binding and replication (Basnet et al., 2019; Everman et al., 2019).

A GWAS of 5 day-use of oral steroids in two pediatric clinical trials revealed a locus in *CTNNA3* that failed to be replicated in an independent study. One SNP at this region was associated with *CTNNA3* gene expression in CD4+ T lymphocytes (McGeachie et al., 2015).

Yan et al. (2021b) conducted a genome-wide association study of asthma hospitalizations in 34,167 white British adults with asthma from the United Kingdom Biobank (UK Biobank). Interestingly, the authors described a novel genetic association for AEs in a gene previously associated with asthma and atopy (Johansson et al., 2019). Specifically, the A allele of rs56151658 *HLA-DQB1* was genome-wide associated with increased odds of hospitalizations for asthma. Additionally, genetic variation in strong LD ($r^2 \geq 0.9$) with rs56151658 regulated the expression of several Human leukocyte antigen class II genes in lung tissue and airway epithelium.

The same authors performed a meta-analysis of GWAS of AEs including 4,010 Hispanic/Latino children and youth with asthma. In this analysis, the SNP rs2253681 (*FLJ22447*) was significantly associated with severe exacerbations at genome-wide level. The risk allele (A) of rs2253681 exerted meQTL effects on a CpG site nearby (cg21115391) by increasing DNA methylation levels in nasal epithelium. In turn, cg21115391 acted as an expression quantitative trait methylation (eQTM) probe that enhanced *KCNJ2-AS1* gene expression (Yan et al., 2021a). Although the specific functions of both long non-coding RNAs remain unknown, *FLI22447* is known to increase interleukin 33 levels and fibroblast expression of several proteins involved in airway remodeling, including vimentin and N-cadherin (Ding et al., 2018).

1.2.6.3 Pharmacogenomic studies

The rest of the GWAS published focused on pharmacogenomics of AEs, including four GWAS in asthma patients under ICS treatment (Dahlin et al., 2015; Hernandez-Pacheco et al., 2019; Dahlin et al., 2020; Hernandez-Pacheco et al., 2021). These have revealed suggestive evidence for association in genes implicated in viral infection (*CMTR1*, Dahlin et al. (2015)) and *APOBEC3B/APOBEC3C*, Hernandez-Pacheco et al. (2019)), baseline lung function (*CACNA2D3*, Hernandez-Pacheco et al. (2021)), bronchodilator responsiveness (*CACNA2D3*, Lutz et al. (2015)), or the Wingless/integrase 1 signaling (*WNT5A*, Hernandez-Pacheco et al. (2021)). Moreover, two genome-wide significant age-by-genotype interactions on AEs despite ICS use were located within genes participating in angiogenesis, lung function and COPD (*THSD4*), and inflammatory, immune, and glucocorticoid response (*HIVEP2*) (Dahlin et al., 2020). Additionally, one study identified two suggestive genetic loci for AEs in asthma patients using long-acting beta2-agonists (Slob et al., 2021) within genes previously involved in lung function (*TBX3*, Soler Artigas et al. (2015)) and response to short beta2-agonists (*EPHA7*, Hardin et al. (2016)).

More recently, the combined analyses including genetic variation and gene expression data have been used to identify genetic variants associated with AEs. One study conducted an eQTL analysis evaluating the role of genetic variation on gene expression of 352 genes differentially expressed in peripheral blood under respiratory syncytial virus infection (Tsai et al., 2020a). One variant within *CEACAM3* regulated the effects of the latent infection on AEs and modified *CEACAM3* gene expression in lung tissue (Tsai et al., 2020a). Moreover, in order to identify relevant genes for a follow-up candidate-gene study, Hernandez-Pacheco et al. (2021) conducted an analysis of differential expression in lung, airway smooth muscle and peripheral blood mononu-

clear in response to ICS treatment. Genetic variation within one of the six differentially expressed genes identified across different tissues (*LTBP1*) was significantly associated with AEs despite ICS use in Europeans and admixed populations at high risk of asthma.

1.2.6.4 Next-generation sequencing-based approaches

None of the genetic studies of AEs published up to date of the writing has implemented next-generation sequencing (NGS) -based approaches to analyse the contribution of common (MAF≥0.05), low-frequency (0.01≤MAF<0.05), or rare (MAF<0.01) variation in AEs either within protein coding, non-protein coding regions of the whole genome. In fact, in those populations that show ethnic disparities on AEs rates, whole-exome or whole-genome sequencing strategies may unveil rare variants with allele frequencies that vary among ethnic backgrounds (Suarez-Pajes et al., 2021). In order to reach sufficient statistical power to detect significant genetic associations, the rare variants are usually grouped by functionally relevant genomic regions (Lin et al., 2021).

1.2.6.5 Admixture mapping

Admixture mapping is a gene mapping strategy applied to admixed populations that show differential levels, profiles or rates of the prevalence of traits of interest. Admixed populations result from the admixture of reproductively isolated ancestral groups; therefore, the genetic background of an admixed individual is a mosaic of chromosomes containing segments of ancestry inherited from the ancestral populations (**Figure 5A**). These differences on the ancestral dosage in each chromosomal region can be leveraged to identify regions of the genome where local ancestry is associated with the trait of interest. These associations usually reflect the presence in the region of one or several variants that confer susceptibility to the trait and usually show differences on the allele frequency among ancestral populations (Shriner, 2017; Hernandez-Pacheco et al., 2016; Suarez-Pajes et al., 2021) (**Figure 5B**).

Admixture mapping can be a powerful alternative to GWAS because the multiple testing correction is based in long-range admixture LD rather than genetic variation LD. Thus, after the genomic region is identified based on the biological hypothesis, the variant or groups of variants responsible for the admixture mapping peaks are identified by fine-mapping these regions (Shriner, 2017; Hernandez-Pacheco et al., 2016; Suarez-Pajes et al., 2021). Admixture mapping can reveal genetic variants that are usually more common in one of the reproductively isolated ancestral groups, followed by the admixed group and less common on other ethnic groups (Shriner, 2017;

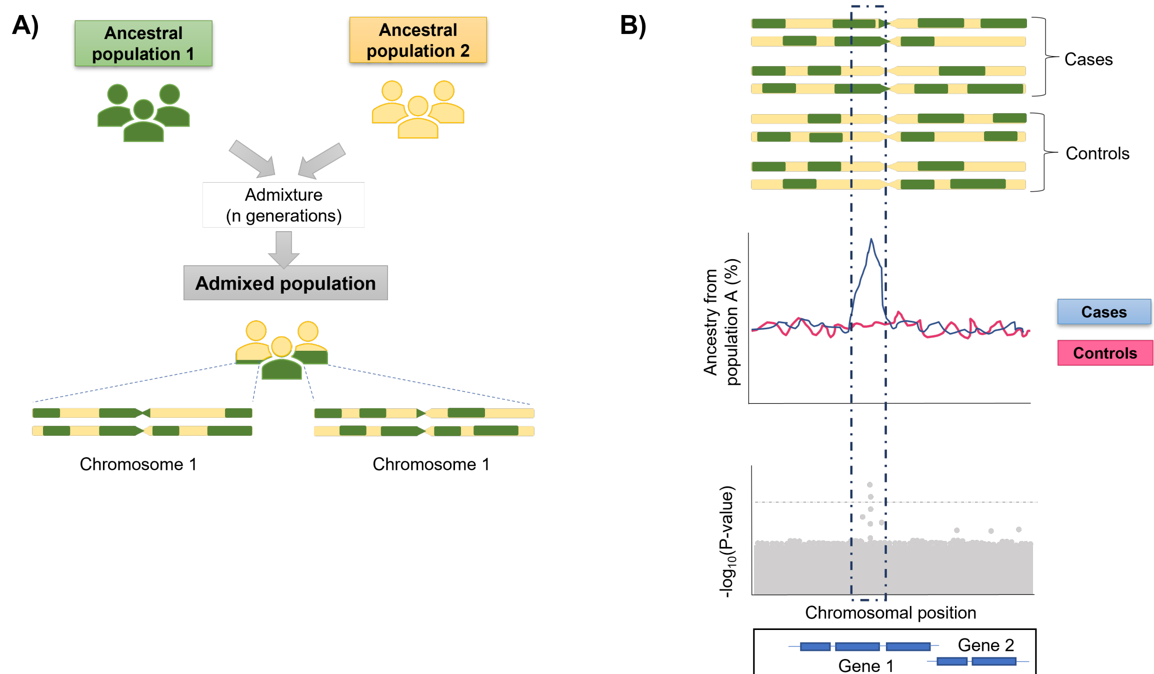


Figure 5. Theoretical framework of admixture mapping. A) Schematic representation of population admixture. Chromosomes of the ancestral populations recombine trough admixture to form the mosaic genomes in the admixed population. **B) Admixture mapping and fine-mapping analyses.** First, in a case-control study, local ancestry estimates are compared to identify genomic regions where local ancestry is associated with the disease status. Then, genetic variants within the significant admixture mapping peak are tested for association with the disease status.

Hernandez-Pacheco et al., 2016; Suarez-Pajes et al., 2021). Moreover, these variant may exert shared-population or population-specific effects that could be explained by differences on allele frequencies or the QTL architecture (Keys et al., 2020). Because of these, NGS techniques and functional prioritization acquire special interest when causal genetic variation is aimed to be uncovered by admixture mapping analyses.

2

Hypotheses and aims

The **general hypothesis** of this work is that AEs are partially influenced by the genetic composition of the individuals. The **specific hypotheses** are:

1. Analyzing populations of several ancestral backgrounds through different genomic approaches may provide complementary information to increase the knowledge of the genetic basis of AEs.
2. Ethnic disparities on AEs rates may reflect ethnic-specific effects and/or differences in the allele frequency of the genetic variants for AEs among populations.

The **main objective** of this doctoral thesis is to attempt to identify genetic loci for AEs and evaluate their functional effects on gene expression and DNA methylation. The **specific aims** are:

1. To review the bibliographic evidence provided by the genetic association studies of AEs to identify potential variants of interest whose association requires validation.
2. To identify genetic variants associated with AEs within genomic regions harbouring genes whose transcript expression levels have been shown to predict AE status.
3. To identify genetic loci associated with the susceptibility to develop asthma with exacerbations in Hispanics/Latinos and African Americans.
4. To identify genetic variation associated with AEs and to validate previous associations across multiple ancestral backgrounds (Europeans, Hispanics/Latinos, African Americans, and Asians).

5. To uncover genomic regions where local genetic ancestry is associated with AEs and to identify potential population-specific genetic variation for this trait in Hispanics/Latinos

6. To evaluate the effects of genetic variation on gene expression and DNA methylation.

3

Materials and methods

3.1 Literature mining of genetic loci for asthma exacerbations

In order to detect genetic variants that could be followed up for replication in independent studies, we conducted a literature search on PubMed (Sayers et al., 2021) of genetic and genomic studies of AEs published up to July 2021. The selected key words were "asthma" AND "exacerbations" OR "hospitalizations" OR "oral corticosteroids" OR "emergency room visits" OR "GWAS" OR "candidate gene" OR "SNP" OR "polymorphism". The genetic associations were classified by study type as follows: candidate-gene approaches, GWAS, gene-environment interactions or pharmacogenetic association.

3.2 Study populations

All *Chapters* in this doctoral thesis comprise data from the Admixture in Latino Americans (GALA II) Study (Borrell et al., 2013; Nishimura et al., 2013; Thakur et al., 2013) and the Study of African Americans, Asthma, Genes, and Environments (SAGE) (Borrell et al., 2013; Nishimura et al., 2013; Thakur et al., 2013). Both studies are parallel case-control studies with similar protocols and questionnaires. The main objective of these studies is to explore the genetic epidemiology of asthma in minority populations at high risk of asthma. While GALA II included children and young adults from five different centers from the mainland United States and Puerto Rico, SAGE recruited individuals from the San Francisco Bay area. Participants were included if they and their ancestors up to the grandparent level self-identified as Hispanic/Latino (GALA II) or African American (SAGE). Inclusion/exclusion criteria have been detailed elsewhere (Borrell et al., 2013; Nishimura et al., 2013; Thakur et al., 2013).

Table 2 summarizes the studies included in each *Chapter* of this doctoral thesis, along with the study design, phenotype definition, and and the array or sequencing method used for the genetic profiling.

Table 2. Overview of the studies included in this doctoral thesis. Acronyms, recruitment country, study design, phenotype definition, genetic profiling platform and inclusion in each *Chapter*.

Study	Study design	Country	Phenotype definition (period of time, months)	Genetic profiling	Chapter
ALLIANCE	Case-control study of asthma	Germany	Hosp/ER/OCS (12)	Infinium Global Screening Array-24 v1.0 +MD BeadChip	5
ALSPAC	Population-based birth cohort	UK	Hosp (12-48)/SA (12)	Illumina HumanHap 550 quad	5
BAMSE	Population-based birth cohort	Sweden	ER/Hosp/SA (12)	Illumina 610 quad, Illumina Infinium Global Screening Array-24	4-5
BREATHE	Asthma cohort	UK	Hosp/OCS/SA (6)	Axiom Precision Medicine Research Array Illumina Infinium CoreExome-24 BeadChip	5
COMPASS	6-month randomized, double-blind, clinical trial	South Africa, Argentina, Bulgaria, Czech Republic, Hungary, Poland, UK Philippines	Hosp/ER/OCS use (12)	Illumina HumanOmniExpress-12v1_C	5
followMAGICs	Asthma cohort	Germany, Austria	Hosp/ER (12)	Illumina Sentrix HumanHap300 BeadChip	5
GALA II	Case-control study of asthma	US, Puerto Rico	Asthma with Hosp/ER/OCS (12)	Affymetrix Axiom LAT1,	3
			Hosp/ER/OCS (12)	Affymetrix Axiom LAT HLA plus array, World Array 4	2, 5
			Hosp/ER/OCS (12)	WGS	3
GEMAS	Asthma cohort	Spain	Hosp/ER/OCS (12)	Illumina Global Screening Array v3.0	4-5
goSHARE	Population-based cohort	Scotland	Hosp/OCS (12)	Axiom Precision Medicine Research Array	5
HPR	Case-control study of asthma	US, Puerto Rico	Hosp/ER/OCS (12)	Illumina 2.5M	5
INMA	Population-based cohort	Spain	Wheezing & asthma (12)	Global Screening Array (several versions), HumanOmni1-Quad v1.0 Beadchip	4
MGI	Population-based cohort	US	Asthma with exacerbation-ICD 495.2 (NA)	Illumina Infinium CoreExome-24 bead	2
MEGA	Case-control study of asthma	Spain	Hosp/ER/OCS (12)	Illumina Global Screening Array v3.0	4-5
PACMAN	Asthma cohort	The Netherlands	Hosp/OCS (12)	Illumina Infinium CoreExome-24 BeadChip	5
PAGES	Asthma cohort	UK	Hosp/OCS (6)	Axiom Precision Medicine Research Array	5
PASS	Asthma cohort	UK	OCS (6)	Illumina Omni Express 8 v1	5
SAGE	Case-control study of asthma	US	Hosp/ER/OCS (12)	Affymetrix Axiom LAT1,	2,5
			Asthma with Hosp/ER/OCS (12)	Affymetrix Axiom LAT HLA plus array, World Array 4	4
				WGS	4
SCAALA	Prospective cohort study of asthma	Brazil	Hosp/ER (12)	Illumina Human Omni2.5–8 v1	5
SCSGES	Asthma cohort	Singapore	Hosp/ER (12)	Infinium OmniZhongHua-8 v1.3 BeadChip, Illumina HumanHap 550 k BeadChip v3, InfiniumOmni2-5 Exome Infinium Global Screening Array	4-5
SLOVENIA	Asthma cohort	Slovenia	Hosp/ER/OCS (12)	Illumina Global Screening Array-24 v1.0 BeadChip	5
The Rotterdam Study	Population-based prospective cohort	The Netherlands	Hosp/ER/OCS (12)	Illumina	5
UBIOPRED	Asthma cohort	Europe	Hosp/ER/OCS (12)	Affymetrix Axiom UK Biobank	5
UK Biobank	Population-based prospective study	UK	Asthma with Hosp (NA)		2
			Hosp (12)	Affymetrix UK Biobank Axiom Array	5

Abbreviations: Hosp: Hospitalizations; ICD: International Classification of Diseases code; ER: Emergency room visits/acute asthma care; OCS: Oral corticosteroids use; SA: School absences; NA: Not available in the source; UK: United Kingdom; US: United States.

In *Chapter 2*, we assessed for validation in Europeans the significant genetic variants using public online databases including data from the UK Biobank and the Michigan Genomics Initiative (MGI) (UK Biobank, 2020; Neale lab, 2018; Michigan Genomics Initiative, 2018) due to the lack of data available from other populations at the time of the analysis.

The multi-ancestry analysis in *Chapter 5* included ancestry-stratified analysis in asthma patients when more than one study from each ethnic group was available. The discovery phase comprised the results from Hispanics/Latinos (GALAII), African Americans (SAGE), and the Singaporean Chinese from the Singapore Cross Sectional Genetic Epidemiology Study (SCSGES) (Andiappan et al., 2010, 2011, 2014; Sio et al., 2020a,b), along with the results from a meta-analysis of GWAS of AEs in European descents. This meta-analysis encompassed GWAS results from participants from the following studies: the Avon Longitudinal Study of Parents and Children (ALSPAC) (Boyd et al., 2013; Fraser et al., 2013; Haag et al., 2020), the BREATHE study (Palmer et al., 2006a,b; Tavendale et al., 2008), the COMPASS study (Kuna et al., 2007), the Genetic of Scottish Health Research Register (GoSHARE) study (McKinstry et al., 2017), the Pharmacogenetics of Asthma Medication in Children: Medication with Anti-inflammatory effects (PACMAN) study (Koster et al., 2009), the Paediatric Asthma Gene-Environment Study (PAGES) (Turner et al., 2010), the Pharmacogenetics of Adrenal Suppression study (PASS) (Hawcutt et al., 2015, 2018), the SLOVENIA study (Berce et al., 2013) and Unbiased Biomarkers for the Prediction of Respiratory Disease Outcomes (U-BIOPRED) study (Shaw et al., 2015). The replication stage comprised data from Europeans from the following studies: the ALL Age Asthma Cohort (ALLIANCE) (Fuchs et al., 2018), Children Allergy Milieu Stockholm an Epidemiological Study (BAMSE) (Kull et al., 2010; Melén et al., 2020), the BREATHE study (Palmer et al., 2006a,b; Tavendale et al., 2008), the follow-up phase of the Multicenter Asthma Genetics in Childhood Study (follow-MAGICS) (Pandey et al., 2013; Schieck et al., 2014; Nieuwenhuis et al., 2016), the Genomics and Metagenomics of Asthma Severity (GEMAS) study (Perez-Garcia et al., 2020), Mechanism underlying the genesis and evolution of asthma (MEGA) project (Muñoz et al., 2018), Rotterdam Study (Ikram et al., 2017)and the UK Biobank. The Hispanic/Latino results from the replication stage included data from participants recruited by the Social Changes, Asthma and Allergy in Latin America (SCAALA) (Barreto et al., 2006) and the Hartford-Puerto Rico (HPR) studies (Yan et al., 2017). Filipino individuals from the COMPASS study (Kuna et al., 2007) with genome-wide genotyping and phenotype data were also included in the replication stage.

The assessment for trans-validations of the genetic variants identified by admixture mapping and subsequent fine-mapping in Hispanic/Latino children and youth with

asthma, as described in *Chapter 4*, included Europeans with asthma from BAMSE, GEMAS, the INfancia y Medio Ambiente (INMA) (Guxens et al., 2012) and MEGA studies along with Singaporean Chinese (SCSGES), Hispanics/Latinos (GALAII), and African Americans (SAGE).

All studies included were approved by their respective institutional review boards and written informed consent was provided by participants or their parents/caregivers. All methods followed the guidelines of the Declaration of Helsinki.

3.3 Trait definition

Health-related outcomes were assessed through questionnaire data and/or quantitative clinical measurements. The definition of the main phenotype of interest, AEs, within each study per *Chapter* is shown in **Table 2**. Apart from the main outcome, other relevant qualitative or quantitative traits were evaluated in secondary analyses: Asthma control (*Chapter 3*), lung function (*Chapter 2-5*), total serum IgE levels (*Chapter 3*), and the presence of asthma with exacerbations and atopic comorbidities (*Chapter 3*).

Asthma control was defined based on five domains evaluated in the Childhood Asthma Control Test (Liu et al., 2007) and the Asthma Control Questionnaire (Juniper et al., 1999). The information about asthma nighttime symptoms, daytime symptoms, activity limitations, rescue medication use, and pulmonary function tests was assessed according to the information available by means of the following questions: a) In general, during the past week how much shortness of breath did the child experience because of asthma?; b) On average during the past week how often was the child woken by his/her asthma during the night?; c) In general, during the past week how limited was the child in his/her activities because of asthma?; d) How many times did the patient report short-acting beta-agonist use in the last two weeks?; e) Percent predicted forced expiratory volume in 1 second and FEV_1/FVC ratio; f) Most recent time the child was prescribed oral steroids for asthma? The first five questions were scored as 0 (best asthma control), 1 (intermediate asthma control), or 2 (worst asthma control). Prescription of oral steroids for asthma was scored as 0 (more than 6 months ago) or 1 (6 months ago or less). The maximum score across the five domains was categorized as very well controlled (0), not well controlled (1), or very poorly controlled (2) asthma.

Pulmonary function testing was conducted with a KoKo® PFTSpirometer (nSpire Health Inc., Louisville, CO) according to ATS recommendations (Crapo et al., 1995).

Percentage predicted lung function values were calculated using the Global Lung Initiative equations (Quanjer et al., 2012). Total serum IgE levels (international units per milliliter) were measured in duplicate on the ImmunoCAP™ 100 system (Phadia, Kalamazooo, MI) and normalized using the common logarithm.

Asthma with severe exacerbations and hay fever/rhinitis or eczema was defined based on the presence of asthma-related hospitalizations, acute care, and/or OCS in the previous last 12 months among individuals with a physician diagnosis of asthma and manifestations of asthma symptoms within the two years before recruitment.

Asthma severity was defined based on medication treatment steps. In *Chapters 2-4*, medication regimen was assessed following the NAEPP 2007 guidelines (U.S. National Heart, Lung, and Blood Institute of the National Institutes of Health, 2007) as follows: Step 1: levels based on the use of short beta-agonists; Step 2: one inhaled corticosteroid, leukotriene inhibitor, or theophylline tablet; and Step 3: more than one or a combination of an inhaled corticosteroid, leukotriene inhibitor, or theophylline tablet or a combination of inhaled corticosteroids and long-beta agonists. Since the analysis described in *Chapter 5* included asthma patients recruited from populations worldwide, we maximized the homogeneity of the criteria using the BTS/SIGN 2014 (British Thoracic Society and Scottish Intercollegiate Guidelines Network, 2014): Step 0: No medication; Step 1: As needed SABA: Step 2: As needed SABA plus regular ICS; Step 2: As needed SABA plus regular ICS and LABA. Step 3: As needed SABA plus regular ICS, LABA, and LTRA. For COMPASS, the medication regimen was adapted from Global Initiative for Asthma and National Asthma Education and Prevention Program guidelines (U.S. National Heart, Lung, and Blood Institute of the National Institutes of Health, 2007). For GALA II and SAGE, the medication regimen was adapted from the British Thoracic Socety/Scottish Intercollegiate Guidelines Network guidelines and Global Initiative for Asthma guidelines (Global Initiative for Asthma, 2021b,a).

3.4 Generation of -omic data, quality control and processing

3.4.1 Genetic variation data

All the studies included in this doctoral thesis had genome-wide genotyping data available (**Table 2**). WGS data from the participants from the GALA II and SAGE studies was made available at the last phase of development of this work, and was used in *Chapter 4*, and assessment of classification models of AEs based on genetic data.

Genetic data was profiled and subjected to quality control following the methods described within the respective original manuscript and/or each *Chapter*.

The ancestral reference panel for the estimation of global and local ancestry described in *Chapters 2-4* comprised 90 HapMap phase II Europeans (CEU) and 90 Yorubas (YRI) with data available from the Axiom Genotype Data Set and 71 Native American (NAM) individuals (17 Zapotec, 2 Mixe, 11 Mixtec from Oaxaca, and 44 Nahua from Central Mexico) genotyped on the Axiom LAT1 array (Pino-Yanes et al., 2015b).

Assessment of the genetic ancestry was conducted using the following procedures, based on the availability of the data: a) Principal component analysis of genome-wide genotyping data (*Chapter 2-3* and *5*), and/or b) Global ancestry calculated from the genomewide-genotyping data with ADMIXTURE (Alexander et al., 2009) (*Chapter 2-3*) or the average of WGS-based local ancestry estimates across the genome (*Chapter 4*).

For *Chapter 4*, samples were phased with BEAGLE 5.0 (Browning and Browning, 2007), and local ancestry was estimated with RFMIX v2.0 (Maples et al., 2013), using phased biallelic SNPs that surpassed the quality control procedures from the WGS data and were available in the ancestral reference panel. A three-way and two-way ancestral models were considered for GALA II (CEU, YRI, and NAM) and SAGE (CEU and YRI), respectively.

In terms of genome-wide genotyping data, the Haplotype Reference Consortium (HRC, r1.1 2016) (McCarthy et al., 2016) was used as reference imputation panel for most of the studies, except for ALSPAC and SCSGES that used the phase 3 of the 1000 Genomes Project (1KGP) and the 1KGP HapMap Han Chinese (CHB) and Chinese (CHD) samples (Consortium et al., 2015), respectively. Moreover, in order to assess the impact of the selected reference imputation panel in *Chapter 3*, we conducted imputation with 1KGP (Consortium et al., 2015) and the Consortium on Asthma among African-ancestry Populations in the Americas (CAAPA) reference (Mathias et al., 2016). Imputation was conducted via Michigan Imputation Server (McCarthy et al., 2016) with Minimac3 (Das et al., 2016) for most of the studies, except for GEMAS, MEGA, and INMA, where Minimac4 was used. In BAMSE and SCGES, imputation was performed with IMPUTE v2.0 (Howie et al., 2009).

3.4.2 DNA methylation data

Bisulfite conversion of 1ug of DNA was performed using the Zymo EZ DNA Methylation Kit (Zymo research, Irvine, CA) following the manufacturer's protocols. For the analyses described in *Chapter 3*, whole-blood DNA methylation from GALA II participants profiled with the Infinium HumanMethylation450 BeadChip array was available. In order to maximize the homogeneity with the whole-blood DNA methylation data from the GALA II and SAGE studies profiled with the Infinium EPIC BeadChip that was made available at the time of the analysis described in *Chapters 4-5*, quality control was re-conducted with *ENmix* (1.22.0) (Xu et al., 2016) R package following the same procedures. Beta values, ranging from 0 to 1, were transformed to M-values as log2(β/(1-β)) for numerical stability (Du et al., 2010). Details on DNA methylation quality control and processing are provided within each *Chapter*. Cell type heterogeneity was captured using the ReFACTor algorithm (Rahmani et al., 2016) within the GLINT 1.0.4 framework (Rahmani et al., 2017), adjusting for age, sex, the top three principal WGS-based components, and batch (serial number as a dummy variable).

3.4.3 Gene expression data

Total RNA was quantified using the Quant-iT™ RiboGreen® RNA Assay Kit and normalized for library preparation with the Illumina TruSeq™ Stranded mRNA Sample Preparation Kit. Libraries were sequenced according to the manufacturer's protocols using the HiSeq 4000 system (Illumina). Quality control and normalization was conducted as described elsewhere (Mak et al., 2021). Additional quality control steps included analysis of sex discordance and the removal of count distribution and population outliers. The *edgeR* R package (Robinson et al., 2010) was used to obtain counts per million and the trimmed mean of M values (TMM). TMM were normalized across samples via the inverse normal transformation. Details on quality control and processing of gene expression data are described in *Chapter 4*. Probabilistic estimation of expression residuals (PEER) was conducted among asthma patients, with adjustment by age, sex and the top three principal WGS-based components (Stegle et al., 2010).

3.5 Statistical and bioinformatic analyses

3.5.1 Admixture mapping analyses

The association of AEs and the number of copies of local ancestry (0, 1, or 2 copies) was tested in R (Team, 2020) through logistic regression models with adjustment by age, sex, global ancestry and asthma severity. These analyses were performed in Puerto Ricans and Mexican Americans from GALA II, separately, as described in *Chapter 4*. To adjust for multiple comparisons, the effective number of tests performed was estimated via an empirical autocorrelation algorithm using the *coda* R package (Plummer et al., 2006), and the statistical threshold of significance was defined as p=0.05/the effective number of tests.

3.5.2 Allelic testing, meta-analysis, and independent variant assessment

Allelic testing was performed separately for each study using EPACTS 3.2.6 (Kang, 2016) (*Chapter 2-4*), PLINK 2.0 (Chang et al., 2015) (*Chapter 4-5*), or rvtests 2.1.0 (Zhan et al., 2016) (*Chapter 5*). The association of genetic variants with qualitative and quantitative traits was assessed using logistic or linear regression models adjusted by potential confounding factors including age, sex, genetic ancestry and/or asthma severity, as detailed within each *Chapter*.

Meta-analyses were conducted with METASOFT (Han and Eskin, 2011) using the inverse-variance-weighted effect size procedure. Fixed-effects or random-effects models were chosen based on the heterogeneity among the studies measured by the Cochran's Q test p-value.

Independent variants were identified by LD pruning using PLINK 1.9 (Chang et al., 2015) in *Chapter 2*, step-wise conditional analysis using GCTA-COJO v1.93.2 (Yang et al., 2011) in *Chapters 3* and *5*, and LD clumping using PLINK 1.9 (Chang et al., 2015) in *Chapter 4*. The references for LD estimation were the imputed dosages or genotypes from SAGE (*Chapter 2*) and GALA II (*Chapter 3*), WGS data from GALA II (*Chapter 4*) or the 1KGP reference (Consortium et al., 2015) (*Chapter 5*).

3.5.3 Genetic associations shared with other traits

In *Chapter 5*, gene-set enrichment analysis in genes reported for different traits in the GWAS catalog (Buniello et al., 2019) was conducted using the GENE2FUNC algorithm, as implemented in the Functional Mapping and Annotation of Genome-Wide Association Studies tool (FUMA GWAS) (Watanabe et al., 2017). For this, SNPs associated at $p \leq 1 \times 10^{-4}$ in the discovery phase were mapped to the closest gene using the UCSC Table Browser tool (Karolchik et al., 2004) to include genes that may be relevant to the trait despite not reaching evidence of suggestive association. A false discovery rate (FDR) of 5% was used to declare significance.

Pairwise genome-wide genetic correlations between AEs and other traits were estimated via comparison with publicly-available GWAS summary statistics using LD score regression, as implemented in LDhub (Zheng et al., 2017). To maximize the homogeneity with publicly-available GWAS summary statistics in Europeans, this analysis was restricted to the results obtained in the European-descent individuals to maximize the statistical power. Traits where the LDHub output included the following warnings were excluded from the analysis: "Caution: using these data may yield less robust results due to minor departure of the LD structure" and "Caution: using this data may yield results outside bounds due to relative low Z score of the SNP heritability of the trait". A Bonferroni-corrected significance threshold was defined p=0.05/number of analyzed traits.

3.5.4 Assessment of functional effects

Quantitative trait loci analyses were carried out using linear regression models via MatrixeQTL (Shabalin, 2012) (*Chapter 3-4*) or fastQTL softwares (*Chapters 4-5*) (Ongen et al., 2016). The regression models were corrected for potential confounders (e.g., age and sex), as detailed within each *Chapter*. Moreover, meQTL models were corrected for cell-type heterogeneity and DNA methylation batch (*Chapter 3-5*), whereas eQTL analyses were adjusted by the top 60 PEER factors (Stegle et al., 2010) (*Chapter 4*), and eQTM analyses were corrected for cell-type heterogeneity, DNA methylation batch, and the top 15 PEER factors (*Chapter 4*). The reduction in the number of PEER factors was motivated by the decreased number of individuals with both DNA methylation and gene expression data available (Stegle et al., 2010). Although in the meQTL analysis described in *Chapter 3*, we evaluated CpG sites within 500 kb of the genetic variant, we followed a more conservative strategy in the following *Chapters* by assessing CpG sites within ±1 Megabase of the genetic variants. *cis*-eQTL analyses considered a region

of ±1 Megabase around the genetic variant, whereas the *cis*-eQTM analysis included genes with the transcription start site located within ±1 Megabase of the significant CpGs. In order to correct for multiple comparisons, we established the significance threshold as a false discovery rate-adjusted p or Storey q < 0.05, as described within each *Chapter*.

In silico evidence of functional effects of the variants on gene expression and DNA methylation was assessed using publicly available data online via QTLbase (Zheng et al., 2020) (*Chapter 5*), the Genotype-Tissue Expression (GTEx) v8 Portal (GTEx Consortium, 2020) (*Chapter 2-5*), the Genetics of DNA Methylation Consortium (Min et al., 2021) (*Chapter 3*), and/or PhenoScanner v2 (Staley et al., 2016; Kamat et al., 2019) (*Chapter 4-5*). Moreover, evidence of DNase I hypersensitive sites and histone modifications marks were evaluated with the Roadmap (Kundaje et al., 2015) and/or ENCODE (Moore et al., 2020) data within HaploReg v4.1 (Ward and Kellis, 2016) (*Chapter 3*), or the SNPnexus framework (Dayem Ullah et al., 2018) (*Chapter 4*). Gene expression data were evaluated with the GTEx Portal (GTEx Consortium, 2020) (*Chapter 2-5*) and RE-ALGAR (Shumyatcher et al., 2017) (*Chapter 4*). Long-distance chromatin interactions were determined using to OpenTarget Genetics (Ghoussaini et al., 2021) (*Chapter 2*) and ChiCP (Schofield et al., 2016) (*Chapter 4-5*). Gene ontology analysis was carried out with OpenTarget Genetics (Ghoussaini et al., 2021) (*Chapter 2-3*).

Annotation of CpG sites was conducted with the manufacturer annotation data via the *IlluminaHumanMethylation450kanno.ilmn12.hg19* R package (Hansen, 2016) (hg19, in *Chapter 3*), the Illumina Human Methylation EPIC annotation file 10b5 (hg38, in *Chapter 4*), the Illumina Human Methylation EPIC annotation file 10b4 (hg19, in *Chapter 5*), or GREAT (McLean et al., 2010) (*Chapter 3-5*). In *Chapter 5*, an enrichment analysis on transcriptional factor binding motifs in lung tissues for the significant CpG sites was performed with eFORGE-TF (Breeze et al., 2019). To correct for multiple testing, the Benjamini-Yekutieli method was used to declare significance (q-value<0.001). To assess enrichment in gene ontology terms for the transcriptional factors that joined to the relevant motifs, gene set enrichment analysis for protein-coding genes was conducted with FUMA GWAS (Watanabe et al., 2017). A Benjamini-Hochberg-based false discovery rate (FDR) of 5% was used to declare significance.

4

CHAPTERS

4.1 Role of genomics in asthma exacerbations

A bibliographic review of all the genetic studies of AEs published up to 15 th December 2018 is reported in this *Chapter*. This work aimed to identify the genetic associations described for AEs, classify them per study type, and analyse the strengths and weaknesses of the published studies in order to prioritize genetic variants to be followed up for replication in independent study populations.

We conducted a search on PubMed using the following key words: “asthma” AND “exacerbations” OR “hospitalizations” OR “oral corticosteroids” OR “emergency room visits” OR “GWAS” OR “candidate gene” OR “SNP” OR “polymorphism”. All references were manually revised by two independent authors and categorized by study type as candidate-gene, GWAS, pharmacogenomics, or gene-environment interaction approaches. Our comprehensive search provided insights into the state of the art of genomic research on AEs performed at the initiation of this thesis, and the future challenges.

This *Chapter* summarizes the findings of an article published in final form in *Current Opinion in Pulmonary Medicine*: Herrera-Luis E, Hernandez-Pacheco N, Vijverberg SJ, Flores C, Pino-Yanes M. Role of genomics in asthma exacerbations. Curr Opin Pulm Med. 2019;25(1):101-112. doi: 10.1097/MCP.0000000000000533. The final peer-reviewed version of the article cannot be reproduced here under the terms of Lippincott Williams & Wilkins license.

Table 3 summarizes the genes implicated in AEs using different approaches. A total of 36 candidate-gene association studies and four genome-wide scans of AEs

had been published up to the stated date, and most of these comprised individuals of European descent. Four studies assessed genetic variation in the *BAFF* (Kumar et al., 2012), *P2RX7* (Denlinger et al., 2013), and *ADRB2* (Ortega et al., 2014; Lima et al., 2006a) genes focused on African American asthma patients, whereas genetic variants in the *TGFB1* (Sharma et al., 2009), *MYLK* (Acosta-Herrera et al., 2015), *SERPINE1* (Cho et al., 2016), *GSTT1* (Turner et al., 2018), *IL10* (Hunninghake et al., 2008), and *IL13* (Hunninghake et al., 2007) genes have been investigated in Hispanics/Latinos. Furthermore, four studies assessed genetic variation in the *SCGB1A1* (Chen et al., 2012), *ADIPOQ* (Ding et al., 2015), *NOS2* (Hirai et al., 2018), and *ADAM33* (Sunadome et al., 2017) genes in Asian populations.

Most of the candidate-gene association studies have been conducted on genes previously implicated in asthma susceptibility or asthma-related traits. In fact, many of these studies have assessed genetic variants included in the 17q21 locus (Blekic et al., 2013; Bisgaard et al., 2009; Tavendale et al., 2008) and the *ADRB2* gene (Turner et al., 2016; Ortega et al., 2014; Zuurhout et al., 2013; Basu et al., 2009; Lima et al., 2006a; Taylor et al., 2000; Palmer et al., 2006b). Moreover, some studies also investigated the role of environmental factors in the modulation of the genetic association for AEs. These environmental factors comprised exposure to dust mite in the *IL9* and *IL10* genes (Sordillo et al., 2015; Hunninghake et al., 2008), fungi in the *CHIT1* gene (Wu et al., 2010), pets in the 17q21 locus (Blekic et al., 2013), endotoxin in the *CD14* and *LY96* genes (Kljaic-Bukvic et al., 2014), or tobacco smoke in the *ADAM33* or 17q21 locus (Blekic et al., 2013; Bukvic et al., 2013), as well as having T2 inflammation in the *ADAM33* and *IL4RA* genes (Sunadome et al., 2017), or respiratory syncytial virus infection in the *SERPINE1* gene (Cho et al., 2016).

Pharmacogenetic candidate-gene studies of AEs have evaluated the role of genetic variation in response to IL-4/IL-13 antagonists, SABA, ICS, LABA, and/or LTRA (**Table 3**). While some studies focused on genes directly involved in drug metabolism, such as *ADRB2* (Turner et al., 2016; Ortega et al., 2014; Zuurhout et al., 2013; Basu et al., 2009; Taylor et al., 2000; Palmer et al., 2006b), *LTA4H*, *LTC4S*, or *ALOX5* (Lima et al., 2006b), others assessed genes involved in key asthma processess, such as regulation of inflammation, including *FCER2* (Koster et al., 2011; Tantisira et al., 2007), *PPARG* (Palmer et al., 2007), or *P2RX7* (Denlinger et al., 2013).

Genome-wide scans of AEs have applied three case-only strategies where asthma patients with and without exacerbations were compared. These have revealed genome-wide significant genetic variants for OCS at the *CTNNA3* locus and suggestive associations in the *SEMA3D* locus (McGeachie et al., 2015), suggestive signals for AEs in response to ICS at the *CMTR1* locus (Dahlin et al., 2015), and genetic variants at *CRTAM*

whose effects on AEs was modulated by vitamin D levels (Du et al., 2012). Moreover, a case-control study including Danish non-asthma subjects and asthma patients with recurrent hospitalizations revealed a genome-wide significant locus at the only known RV-C receptor (*CDHR3*) (Bønnelykke et al., 2014).

A plethora of definitions of AEs, including both quantitative and qualitative variables, have been used in the genetic studies of AEs (**Table 3**). Despite this, the most common definition includes hospitalizations, emergency room visits, and/or OCS use. Besides, the sample size of these studies is limited (from 61 to 3,684 individuals), although there is a trend towards the increase in sample size in recent years, which could improve the power to detect associations in future studies. Furthermore, few of these studies provided evidence of the functional role of the genetic signals, which reduces the understanding of the biological mechanisms involved in AEs. Likewise, the variants revealed to date exhibit moderate effect sizes, which limits their clinical applicability.

Table 3. Summary of genes where genetic variation has been implicated in AEs.

Gene	Genetic variants	Phenotype	Study type†	Reference
CHI3L1	rs4950928	Hosp	CG	Cunningham et al. (2011)
CHIT1	rs1417149	Hosp/ER	CG (Fungi)	Wu et al. (2010)
IL10	rs1800896, rs3024496	Hosp	CG (Dust mite)	Hunninghake et al. (2008)
IL1R1	rs1558641	Asthma with recurrent Hosp	GWAS	Bønnelykke et al. (2014)
PPARG	rs1801282, rs3856806	Hosp/school absence/OCS	CG (LTRA)	Palmer et al. (2007)
MYLK	rs77820417	Hosp/ER/OCS	CG	Acosta-Herrera et al. (2015)
ADIPOQ	rs1501299	Recurrent wheezing/dyspnea/chest tightness/cough concurrent with a positive bronchodilator test	CG	Palmer et al. (2007)
RAD50	rs6871536	Asthma with recurrent Hosp	GWAS	Bønnelykke et al. (2014)
IL13	rs1800925	Hosp/ER	CG (ICS)	Hunninghake et al. (2007)
IL9	rs11741137, rs2069885	Hosp/ER	CG (Dust mite)	Sordillo et al. (2015)
CD14	rs574445563 rs29158	Hosp Hosp	CG CGI (Endotoxin)	Kljaic-Bukvic et al. (2014)
ADRB2	rs1042718	Hosp/ER/OCS/higher ICS dose/higher rescue use/OH/FEV_1/PEF/systemic corticosteroids	CG	Lima et al. (2006a)
ADRB2	rs1042713	Rescue medication/PEF/asthma symptoms	CG (SABA)	Taylor et al. (2000)

†For candidate-gene interaction studies, the exposure is specified in parenthesis. Pharmacogenetic studies show the medication considered in parenthesis. Abbreviations: CG: candidate-gene association study; CGI: Candidate-gene interaction study; ER, emergency room; FEV_1, forced expiratory volume in 1 second; GWIS: genome-wide interaction scan; Hosp, hospitalization; ICU, intensive care unit; I/D: Insertion/deletion; OCS, systemic corticosteroids; ICS: inhaled corticosteroids; LABA: long-acting beta2 agonists; LTRA: leukotriene receptor antagonists; OH: outpatient healthcare; PEF, peak expiratory flow; RSV: respiratory syncytial virus; SABA: short-acting beta2 agonists.

Table 3 (continuation). Summary of genes where genetic variation has been implicated in AEs.

Gene	Genetic variants	Phenotype	Study type[†]	Reference
ADRB2	rs1042713	Hosp/school absence/OCS	CG (SABA/LABA) CG (LABA)	Basu et al. (2009) Palmer et al. (2006)
		ER/OCS	CG (ICS+LABA)	Zuurhout et al. (2013)
		Hosp/ER/OCS/School absence	CG (ICS+LABA)	Turner et al. (2016)
	rs1800888, -376 I/D	OH/ER/Hosp/OCS	CG (LABA)	Ortega et al. (2014)
LTC4S	rs730012	Hosp/ER/OCS/higherICSdose/FEV_1/OH	CG (LTRA)	Lima et al. (2006a)
CMTR1	rs2395672	Hosp/ER/OCS	GWAS (ICS)	Dahlin et al. (2015)
SEMA3D	rs993312	OCS	GWAS	McGeachie et al. (2015)
SERPINE1	rs2227631	Hosp Hosp/ER	CG CGI (RSV)	Cho et al. (2016)
CDHR3	rs6967330	ER Asthma with recurrent Hosp	CG GWAS	Stenberg et al. (2018) Bønnelykke et al. (2014)
LY96	rs17226566	Hosp	CG CGI (Endotoxin)	Kljaic-Bukvic et al. (2014)
IL33	rs928413	Asthma with recurrent Hosp	GWAS	Bønnelykke et al. (2014)
ALOX5	Repeat variant	Hosp/ER/OCS/higher ICS dose/FEV_1/OH	CG (LTRA)	Lima et al. (2006a)
CTNNA3	rs7915695	OCS	GWAS	McGeachie et al. (2015)
SCGB1A1	rs3741240	Hosp/ER/FEV_1/OH/PEF/OCS/higher rescue use	CG	Chen et al. (2012)

[†]For candidate-gene interaction studies, the exposure is specified in parenthesis. Pharmacogenetic studies show the medication considered in parenthesis. Abbreviations: CG: candidate-gene association study; CGI: Candidate-gene interaction study; ER, emergency room; FEV_1, forced expiratory volume in 1 second; GWIS: genome-wide interaction scan; Hosp, hospitalization; ICU, intensive care unit; I/D: Insertion/deletion; OCS, systemic corticosteroids; ICS: inhaled corticosteroids; LABA: long-acting beta2 agonists; LTRA: leukotriene receptor antagonists; OH: outpatient healthcare; PEF, peak expiratory flow; RSV: respiratory syncytial virus; SABA: short-acting beta2 agonists.

Table 3 (continuation). Summary of genes where genetic variation has been implicated in AEs.

Gene	Genetic variants	Phenotype	Study type†	Reference
CRTAM	rs2272094	Hosp/ER	GWIS (Vitamin D levels)	Du et al. (2012)
LTA4H	rs2660845	Hosp/ER/OCS/higher ICS dose/FEV_1/OH	CG (LTRA)	Lima et al. (2006a)
P2RX7	rs2230911	OCS	CG (ICS + LABA)	Denlinger et al. 2013
BAFF	rs17564816	Hosp/ER/OCS	CG (ICS)	Kumar et al. 2012
IL4RA	rs1801275	Mechanical ventilation/ICU care	CG	Wenzel et al. 2007
	rs8832	OCS	CGI (T2 inflammation)	Sunadome et al. 2017
		Hosp/ER/OCS/	CG (Pitrakinra)	Slager et al. 2012
NOS2	rs869282985	OCS	CG	Hirai et al. 2018
ORMDL3	rs3744246	Hosp	CGI (Pets)	Blekic et al. (2013)
	rs8079416		CGI (Early-life exposure to tobacco)	
GSDMB	rs7216389	Hosp/OCS/high dose ICS	CG	Bisgaard et al. (2009)
		Hosp/school absence/OCS	CG	Tavendale et al. (2008)
	rs7212938	Hosp	CGI (Early-life exposure to tobacco)	Blekic et al. (2013)
	rs2305480	Asthma with recurrent Hosp	GWAS	Bønnelykke et al. (2014)
FCER2	rs28364072	Hosp/ER	CG (ICS)	Koster et al. (2011),Tantisira et al. (2007)

†For candidate-gene interaction studies, the exposure is specified in parenthesis. Pharmacogenetic studies show the medication considered in parenthesis. Abbreviations: CG: candidate-gene association study; CGI: Candidate-gene interaction study; ER, emergency room; FEV_1, forced expiratory volume in 1 second; GWIS: genome-wide interaction scan; Hosp, hospitalization; ICU, intensive care unit; I/D: Insertion/deletion; OCS, systemic corticosteroids; ICS: inhaled corticosteroids; LABA: long-acting beta2 agonists; LTRA: leukotriene receptor antagonists; OH: outpatient healthcare; PEF, peak expiratory flow; RSV: respiratory syncytial virus; SABA: short-acting beta2 agonists.

Table 3 (continuation). Summary of genes where genetic variation has been implicated in AEs.

Gene	Genetic variants	Phenotype	Study type†	Reference
TGFB1	rs1800470	Hosp/ER	CG CGI (Dust mite)	Sharma et al. (2009)
GSDMB	rs7216389	Hosp/OCS/high dose ICS	CG	Bisgaard et al. (2009)
		Hosp/school absence/OCS	CG	Tavendale et al. (2008)
	rs7212938	Hosp	CGI (Early-life exposure to tobacco)	Blekic et al. (2013)
	rs2305480	Asthma with recurrent Hosp	GWAS	Bønnelykke et al. (2014)
FCER2	rs28364072	Hosp/ER	CG (ICS)	Koster et al. (2011),Tantisira et al. (2007)
TGFB1	rs1800470	Hosp/ER	CG CGI (Dust mite)	Sharma et al. (2009)
ADAM33	rs512625	Hosp	CG CGI (Tobacco smoke)	Kljaic-Bukvic et al. (2013)
	rs2280090	OCS	CGI (T2 inflammation)	Sunadome et al. (2017)
SIGLEC1	rs532448	Hosp	CG	Kljaic-Bukvic et al. (2013)
GSTT1	*Null/Active* genotype	Hosp/ER/OCS	CG	Turner et al. (2018)
ST13	rs138335	Hosp/ER	CG (ICS)	Vijverberg et al. (2015)

†For candidate-gene interaction studies, the exposure is specified in parenthesis. Pharmacogenetic studies show the medication considered in parenthesis. Abbreviations: CG: candidate-gene association study; CGI: Candidate-gene interaction study; ER, emergency room; FEV_1, forced expiratory volume in 1 second; GWIS: genome-wide interaction scan; Hosp, hospitalization; ICU, intensive care unit; I/D: Insertion/deletion; OCS, systemic corticosteroids; ICS: inhaled corticosteroids; LABA: long-acting beta2 agonists; LTRA: leukotriene receptor antagonists; OH: outpatient healthcare; PEF, peak expiratory flow; RSV: respiratory syncytial virus; SABA: short-acting beta2 agonists.

4.2 A Deoxyribonuclease 1 Like 3 genetic variant associates with asthma exacerbations

This *Chapter* describes the results of a candidate-gene study of severe AEs conducted on six genes, whose combined sputum gene expression signature can discriminate inflammatory phenotypes and predict future AEs. This six-gene signature comprised the expression levels of the following genes: alkaline phosphatase (*ALPL*), the Charcot-Leyden crystal galectin (*CLC*), the carboxypeptidase A3 (*CPA3*), the C-X-C motif chemokine receptor 2 (*CXCR2*), the deoxyribonuclease 1 like 3 (*DNASE1L3*) and the interleukin 1 beta (*IL1B*)). In this study, we tested the association of severe AEs and genetic variants within the genomic regions where where those six genes were located. The region-level significance thresholds were established based on the number of independent genetic variants within each genomic region.

In the discovery stage, comprised of African Americans, the SNP rs67622929 in the *DNASE1L3* region showed significant association with severe AEs after multiple testing correction. None of the polymorphisms within the five remaining genomic regions surpassed the significance thresholds in the discovery stage. The significant genetic variant was followed up for replication in Hispanic/Latinos and showed a significant and consistent direction of the effect in both ethnic groups ($p<0.05$). Moreover, *in silico* assessment of functional effects revealed that the risk allele of rs67622929 increases *DNASE1L3* expression in several tissues. In conclusion, our analysis identified a susceptibility locus for AEs in African Americans and Latinos with regulatory effects on *DNASE1L3* expression.

This *Chapter* was published as a Letter to the Editor entitled 'A Deoxyribonuclease 1 Like 3 genetic variant associates with asthma exacerbations', contributor Maria Pino-Yanes [Herrera-Luis E, Lorenzo-Diaz F, Samedy-Bates LA, Eng C, Villar J, Rodriguez-Santana JR, Burchard EG, Pino-Yanes M.] in the *Journal of Allergy and Clinical Immunology*, volume 147, and issue 3 in 2021 (doi: 10.1016/j.jaci.2020.09.027). Copyright © 2020 Elsevier B.V. Published by Elsevier B.V. This article is reproduced under the terms of Elsevier B.V. license.

2. Wang AL, Datta S, Weiss ST, Tantisira KG. Remission of persistent childhood asthma: early predictors of adult outcomes. J Allergy Clin Immunol 2019;143:1752-9.e6.
3. McGeachie MJ, Davis JS, Kho AT, Dahlin A, Sordillo JE, Sun M, et al. Asthma remission: predicting future airways responsiveness using an miRNA network. J Allergy Clin Immunol 2017;140:598-600.e8.
4. The Childhood Asthma Management Program (CAMP): design, rationale, and methods. Childhood Asthma Management Program Research Group. Control Clin Trials 1999;20:91-120.
5. Li J, Panganiban R, Kho AT, McGeachie MJ, Farnam L, Chase RP, et al. Circulating microRNAs and treatment response in childhood asthma. Am J Respir Crit Care Med 2020;202:65-72.
6. Sears MR. Predicting asthma outcomes. J Allergy Clin Immunol 2015;136:829-36.
7. Perry MM, Baker JE, Gibeon DS, Adcock IM, Chung KF. Airway smooth muscle hyperproliferation is regulated by microRNA-221 in severe asthma. Am J Respir Cell Mol Biol 2014;50:7-17.
8. Xu H, Gu LN, Yang QY, Zhao DY, Liu F. MiR-221 promotes IgE-mediated activation of mast cells degranulation by PI3K/Akt/PLCγ/Ca(2+) pathway. J Bioenerg Biomembr 2016;48:293-9.
9. Fry RC, Rager JE, Bauer R, Sebastian E, Peden DB, Jaspers I, et al. Air toxics and epigenetic effects: ozone altered microRNAs in the sputum of human subjects. Am J Physiol Lung Cell Mol Physiol 2014;306:L1129-37.

Available online Sept 2, 2020.
https://doi.org/10.1016/j.jaci.2020.08.019

A deoxyribonuclease 1–like 3 genetic variant associates with asthma exacerbations

To the Editor:

Recently, Fricker et al[1] reported that an expression signature in the sputum efficiently predicted exacerbations among patients with asthma, showing the greatest performance for frequent exacerbators. This signature was based on 6 genes (*ALPL, CLC, CPA3, CXCR2, DNASE1L3*, and *IL1B*) that can differentiate airway asthmatic inflammatory phenotypes. Thus, we hypothesized that genetic variation within those genes may contribute to susceptibility to exacerbations. In the United States, minority populations of African descent show the highest rates of asthma exacerbations. In fact, African ancestry is a risk factor not only for asthma in Puerto Ricans and African Americans but also for exacerbations among African Americans.[2,3] Hence, we performed a candidate gene association study of asthma exacerbations in African Americans and Latinos.

Details on the study populations and methods are described in the Online Repository. Genotype dosages that have an imputation quality (R^2) of 0.3 or more and a minor allele frequency of 5% or more and are located within the gene limits plus or minus 100 kilobases were examined for association with asthma exacerbations. In the discovery, 1002 African American patients with asthma from the Study of African Americans, Asthma, Genes, and Environments (SAGE), 470 of whom experienced exacerbations, were analyzed (Table I). Asthma exacerbations were defined as a dichotomous variable (presence or absence) based on oral corticosteroid use, hospitalizations, or acute care due to asthma in the 12 months before study enrollment. Association testing was performed by using logistic regression models with correction for age, sex, and 2 principal components (PCs). To account for multiple comparisons, we established a Bonferroni-corrected threshold for significance based on the number of independent single-nucleotide polymorphisms (SNPs) within each genomic region by linkage disequilibrium pruning ($r^2 < 0.2$). Replication was evaluated in 2181 Latinos with asthma from the Genes-Environments and Admixture in Latino Americans (GALA II) study, including 1283 exacerbators (Table I). This 2-stage design provided greater than 80% statistical power to detect the association of common variants for different scenarios (see Table E1 in this article's Online Repository at www.jacionline.org).

A total of 4089 SNPs were evaluated (see Table E2 in this article's Online Repository at www.jacionline.org). Among African Americans, an intronic variant in the deoxyribonuclease 1–like 3 gene (*DNASE1L3*) region (Fig 1), rs67622929, exceeded the significance threshold corrected by multiple comparisons (0.05/57 independent variants [$P = 8.77 \times 10^{-4}$]) (odds ratio [OR] for the C allele = 1.48 [95% CI = 1.18-1.87]; $P = 7.9 \times 10^{-4}$; imputation $R^2 = 0.99$). The association of rs67622929 was replicated in Latinos (OR = 1.18 [95% CI = 1.01-1.37]; $P = .03$; imputation $R^2 = 0.97$). An inverse-variance weighted fixed-effect meta-analysis of effects was performed because no heterogeneity was detected (Cochran Q $P = .103$), revealing that the C allele was significantly associated with increased odds of exacerbations (OR = 1.26 [95% CI 1.11-1.43]; $P = 2.6 \times 10^{-4}$). Population stratification adjustment by African ancestry in African Americans and African and Native ancestry in Latinos, instead of PCs, revealed similar results (see Table E3 in this article's Online Repository at www.jacionline.org). Furthermore, the C allele of rs67622929 was also associated with increased risk of exacerbations when adjusted by disease severity in a subset of African Americans with data available (OR = 1.40 [95% CI = 1.10-1.78]; $P = 6.6 \times 10^{-3}$) and Latinos (OR = 1.18 [95% CI 1.01-1.38]; $P = .034$), as well as in the meta-analysis (OR = 1.24 [95% CI = 1.09-1.41]; $P = 1.2 \times 10^{-3}$; Cochran Q $P = .250$). To explore whether this association was population specific, the SNP rs67622929 was evaluated for validation in individuals of European descent through online databases, and no evidence of association was found ($P > .05$) (see Table E4 in this article's Online Repository at www.jacionline.org). However, the definition of the control group differed from that in our study, given that individuals without asthma were analyzed instead of patients with asthma who were nonexacerbators. Therefore, the association of this SNP with severe asthma exacerbations among populations of European descent cannot be fully ruled out.

Next, we assessed the association of rs67622929 with asthma-related traits, including IgE levels and normalized lung function values at baseline. The linear regression models for normalized lung function values were adjusted by the 2 PCs. For $\log_{10}$ IgE levels, age and sex were also included as covariates. The SNP rs67622929 was not associated ($P > .05$) with FEV_1, forced vital capacity, ratio of FEV_1 to forced vital capacity, or IgE levels (see Table E3).

According to the Genotype-Tissue Expression (GTEx) database,[4] the C allele of rs67622929 acts as an expression quantitative trait locus that increases the expression of *DNASE1L3* in several tissues, including cultured fibroblasts ($p_{GTEX} = 7.1 \times 10^{-16}$) (see Fig E1, *A* and Table E5 in this article's Online Repository at www.jacionline.org). This SNP is also a *cis*-splicing quantitative trait locus for *DNASE3L1* in the lung ($p_{GTEX} = 9.4 \times 10^{-9}$) (see Fig E1, *B*) and other tissues (see Table E6 in this article's Online Repository at www.jacionline.org). Furthermore, rs67622929 has been linked to *FLNB* in naive CD8 cells trough promoter capture Hi-C, according to OpenTargets.[5] *FLNB* is located close to *DNASE1L3* (Fig 1), and

1096 LETTERS TO THE EDITOR

J ALLERGY CLIN IMMUNOL
MARCH 2021

TABLE I. Characteristics of the subjects included in the analyses

	SAGE (N = 1002)		GALA II (N = 2181)	
Characteristic	**Exacerbators (n = 470)**	**Nonexacerbators (n = 532)**	**Exacerbators (n = 1283)**	**Nonexacerbators (n = 898)**
Sex (% female)	229 (48.7)	286 (50.4)	582 (45.4)	403 (44.9)
Age (y), mean ± SD	14.3 ± 5.6*	14.7 ± 4.1	12.2 ± 3.1*	13.4 ± 3.5
African ancestry (%)	79.4 ± 10.7	78.4 ± 12.1	18.9 ± 13.2*	12.6 ± 11.9
Native American ancestry (%)	NA	NA	22.9 ± 22.6*	37.4 ± 25.5
Asthma control, n (%)				
Well controlled	60 (15.7)	182 (36.8)	210 (16.4)*	431 (48.0)
Partially controlled	112 (29.4)	150 (30.4)	533 (41.5)*	282 (31.4)
Poorly controlled	209 (54.9)	162 (32.8)	540 (42.1)*	185 (20.6)
Treatment category, n (%)				
Step 1	111 (23.6)	249 (47.1)	210 (16.4)*	431 (48.0)
Step 2	183 (39.0)	210 (39.7)	533 (41.5)*	282 (31.4)
Step 3	176 (37.4)	70 (13.2)	540 (42.1)*	185 (20.6)
Baseline lung function				
FEV_1 (% predicted), mean ± SD	91.2 ± 11.9	92.4 ± 11.6	97.9 ± 15.7*	103.6 ± 14.4
FVC (% predicted), mean ± SD	96.5 ± 11.8	97.0 ± 10.9	102.0 ± 16.0*	107.3 ± 15.3
FEV_1/FVC (% predicted), mean ± SD	94.3 ± 8.3	95.3 ± 7.6	96.4 ± 7.6*	97.2 ± 7.2
Total serum IgE level (IU/mL), mean ± SD	168.3 ± 4.9*	139.7 ± 4.6	254.0 ± 4.5*	147.4 ± 4.7

FVC, Forced vital capacity; *NA*, not available/applicable.
The severity medication regime was categorized into 3 levels based on the use of short-acting β-agonists (step 1); 1 inhaled corticosteroid, leukotriene inhibitor, or theophylline tablet (step 2); more than 1 or a combination of an inhaled corticosteroid, leukotriene inhibitor, or theophylline tablet or a combination of inhaled corticosteroids and long-acting β-agonists (step 3). Asthma control and treatment category were available for 875 and 999 individuals from SAGE, respectively. Asthma control and severity were available for 2181 and 2170 individuals from GALA II, respectively. FEV_1 value was available for 807 and 2068 SAGE and GALA II subjects. FVC value was available for 821 participants from SAGE and 2077 subjects from GALA II. FEV_1/FVC ratio was available for 804 and 2068 SAGE and GALA II subjects. Total serum IgE levels were available for 784 and 1856 SAGE and GALA II subjects.
*$P < .01$ for the comparison between asthma cases with exacerbations and controls.

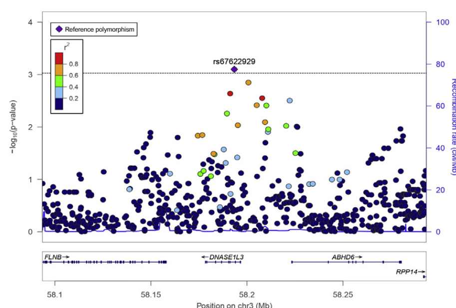


FIG 1. Regional plot for *DNASE1L3* in the discovery sample. The top hit (rs67622929) is shown as a purple diamond, and the rest of the genetic variants are colored on the basis of their linkage disequilibrium with it by using pairwise r^2 values from African populations from the 1000 Genomes Project. Chromosome (Chr) positions, the $-\log_{10}$-transformed *P* values, and the recombination rate are plotted in the x-axis and right and left y-axes, respectively. The Bonferroni-corrected threshold for significance is indicated by the dashed line.

it is transcribed in the opposite direction. *FLNB* is involved in actin, cadherin, and RNA binding, as well as in many biologic processes, including antiviral mechanisms, cellular response to IFN-γ, epithelial cell morphogenesis, and actin cytoskeleton organization. Interestingly, Flnb-deficient mice show skeletal abnormalities and decreased cellular adhesion.[5]

DNASE1L3 encodes an endonuclease that participates in apoptosis, clearing DNA from circulating apoptotic bodies, and it has an important role in autoimmunity protection.[6] Additionally, DNASE1L3 participates in different catabolic and metabolic processes, neutrophil activation, cell development, regulation of inflammasome activation, and cytokine secretion.[5,6] Genetic variants from *DNASE1L3* have previously been associated with neutrophil counts and systemic rheumatic diseases[5,6]; however, this is the first report of an association with asthma exacerbations. Interestingly, the allele associated with risk for asthma exacerbations in our study is associated with higher *DNASE1L3* expression, an effect that is in the same direction as the one described for this gene in the expression signature associated with asthma exacerbations in the previous study.[1] Additionally, *DNASE1L3* is overexpressed in induced sputum from patients with eosinophilic asthma[1,7] and in endobronchial samples from patients with asthma compared with samples from healthy controls.[8] Despite all this evidence, the mechanistic implication of *DNASE1L3* in asthma remains unclear, and functional studies are needed to disentangle it.

We acknowledge some limitations of our study. First, the association of rs67622929 would not be considered significant after application of a more restrictive correction for multiple comparisons accounting for all polymorphisms tested across the 6 genomic regions (0.05/384 independent variants = 1.30×10^{-4}). Second, the sample size of this study is modest compared with those of genetic studies of asthma. To our knowledge, however, this is one of the largest candidate gene association studies of severe asthma exacerbations published to date. Third, heterogeneity in genetic admixture exists in the populations analyzed. Although African Americans are descendants of African and European ancestors, Latinos also have Native American admixture, with varying proportions of each component depending on the Latino subgroup.[9] In this regard,

ethnicity-stratified analyses revealed that the effect in Latinos was driven mainly by Mexicans and other Latinos (see Table E7 in this article's Online Repository at www.jacionline.org).

In conclusion, the SNP rs67622929 was associated with severe asthma exacerbations in African Americans and Latinos and showed regulatory effects over *DNASE1L3* expression.

The authors thank the patients, families, recruiters, health care providers, community clinics, and the study coordinator (Sandra Salazar) participating in GALA II and SAGE.

Esther Herrera-Luis, MSc[a]
Fabian Lorenzo-Diaz, PhD[a]
Lesly-Anne Samedy-Bates, PharmD, PhD[b,c]
Celeste Eng, BSc[b]
Jesús Villar, MD[d,e,f]
Jose R. Rodriguez-Santana, MD[g]
Esteban G. Burchard, MD, MPH[b,c]
Maria Pino-Yanes, PhD[a,d,h]

From [a]the Genomics and Health Group, Department of Biochemistry, Microbiology, Cell Biology and Genetics, Universidad de La Laguna, San Cristóbal de La Laguna, Santa Cruz de Tenerife, Spain; [b]the Department of Medicine, University of California San Francisco, San Francisco, Calif; [c]the Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, Calif; [d]the Centro de Investigación Biomédica en Red de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain; [e]the Multidisciplinary Organ Dysfunction Evaluation Research Network, Research Unit, Hospital Universitario Dr Negrín, Las Palmas de Gran Canaria, Spain; [f]the Keenan Research Center for Biomedical Science at the Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Ontario, Canada; [g]the Centro de Neumología Pediátrica, San Juan, Puerto Rico; and [h]the Instituto de Tecnologías Biomédicas, Universidad de La Laguna, San Cristóbal de La Laguna, Santa Cruz de Tenerife, Spain. E-mail: mdelpino@ull.edu.es.

This study was funded by the Spanish Ministry of Science, Innovation and Universities (SAF2017-83417R), the State Research Agency, and the European Regional Development Fund from the European Union (MICIU/AEI/FEDER, UE). E.H.-L. was funded by a fellowship (PRE2018-083837) from the Spanish Ministry of Science, Innovation, and Universities. M.P.-Y. was supported by the Ramón y Cajal Program by the Spanish Ministry of Science, Innovation, and Universities (RYC-2015-17205). The Genes-Environments and Admixture in Latino Americans (GALA II) Study and the Study of African Americans, Asthma, Genes and Environments (SAGE) were supported in part by the Sanler Family Foundation, the American Asthma Foundation, the RWJF Amos Medical Faculty Development Program, the Harry Wm. and Diana V. Hind Distinguished Professor in Pharmaceutical Sciences II award, the National Institutes of Health, National Heart, Lung, and Blood Institute (R01HL117004, R01HL128439, R01HL135156, 1X01HL134589, R01HL141992, and R01HL141845), the National Institute of Health and Environmental Health Sciences (R01ES015794 and R21ES24844), the National Institute on Minority Health and Health Disparities (P60MD006902, RL5GM118984, R01MD010443, and R56MD013312), the Tobacco-Related Disease Research Program (under awards 24RT-0025 and 27IR-0030), and the National Human Genome Research Institute (U01HG009080). E.H.-L. reports a fellowship from the Spanish Ministry of Science, Innovation, and Universities.

Disclosure of potential conflict of interest: The authors declare that they have no relevant conflicts of interest.

REFERENCES

1. Fricker M, Gibson PG, Powell H, Simpson JL, Yang IA, Upham JW, et al. A sputum 6-gene signature predicts future exacerbations of poorly controlled asthma. J Allergy Clin Immunol 2019;144:51-60.e11.
2. Grossman NL, Ortega VE, King TS, Bleecker ER, Ampleford EA, Bacharier LB, et al. Exacerbation-prone asthma in the context of race and ancestry in Asthma Clinical Research Network trials. J Allergy Clin Immunol 2019;144:1524-33.
3. Hernandez-Pacheco N, Flores C, Oh SS, Burchard EG, Pino-Yanes M. What ancestry can tell us about the genetic origins of inter-ethnic differences in asthma expression. Curr Allergy Asthma Rep 2016;16:53.
4. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580-5.
5. Koscielny G, An P, Carvalho-Silva D, Cham JA, Fumis L, Gasparyan R, et al. Open Targets: a platform for therapeutic target identification and validation. Nucleic Acids Res 2017;45:D985-94.
6. Shi G, Abbott KN, Wu W, Salter RD, Keyel PA. Dnase1L3 regulates inflammasome-dependent cytokine secretion. Front Immunol 2017;8:522.
7. Baines KJ, Simpson JL, Wood LG, Scott RJ, Fibbens NL, Powell H, et al. Sputum gene expression signature of 6 biomarkers discriminates asthma inflammatory phenotypes. J Allergy Clin Immunol 2014;133:997-1007.
8. Sánchez-Ovando S, Baines KJ, Barker D, Wark PA, Simpson JL. Six gene and TH2 signature expression in endobronchial biopsies of participants with asthma. Immunity, Inflamm Dis 2020;8:40-9.
9. Bryc K, Durand EY, Macpherson JM, Reich D, Mountain JL. The genetic ancestry of African Americans, Latinos, and European Americans across the United States. Am J Hum Genet 2015;96:37-53.

Available online Oct 6, 2020.
https://doi.org/10.1016/j.jaci.2020.09.027

Nasopharyngeal *Haemophilus* and local immune response during infant respiratory syncytial virus infection

To the Editor:

The accumulating evidence suggests that viral-bacterial interactions can affect short- and long-term outcomes of acute respiratory infections (ARIs) due to respiratory syncytial virus (RSV) in infancy. In particular, prior studies have found that in young children with RSV ARIs, a higher relative abundance of *Haemophilus* in the nasopharynx is associated with an increased viral load,[1] delayed viral clearance,[2] a different gene expression profile,[3,4] and more severe disease.[3,4] However, the mechanisms underlying these associations are largely unknown. To address this gap in knowledge, we examined the association of the nasopharyngeal relative abundance of *Haemophilus* with viral load and 52 local immune mediators in 105 infants with an RSV-only ARI (ie, no coinfections) who were enrolled in the Infant Susceptibility to Pulmonary Infections and Asthma following RSV Infection in Infancy (INSPIRE) study.

The INSPIRE cohort is a population-based, birth cohort of healthy, term infants born in middle Tennessee with biweekly surveillance of acute respiratory symptoms during their first winter viral season to capture their initial RSV ARI. Infants who met prespecified criteria for an ARI underwent an in-person visit, which included collection of a nasal wash that was used for RSV detection and viral load assessment by real-time RT-PCR. This nasal wash has also been used for characterization of the nasopharyngeal microbiome and local immune response through 16S ribosomal RNA (rRNA) sequencing and measurement of multiple local immune mediators (ie, cytokines, chemokines, and growth factors) by using Luminex xMAP technology (see Table E1 in the Online Repository at www.jacionline.org), respectively, as previously described.[5,6] In brief, following bacterial DNA extraction, we amplified the V4 region of the 16S rRNA gene by using universal primers. The libraries were then sequenced on an Illumina MiSeq platform with 2 $\times$ 300-bp reads. We subsequently processed the 16S rRNA paired sequences by using the dada2 pipeline, after which we grouped sequences into amplicon sequence variants (ASVs) and assigned taxonomy using the SILVA reference database. We used the relative abundance of *Haemophilus* and *Haemophilus* ASVs (expressed as simple proportions) for statistical analyses. In addition, we used the Basic Local Alignment Search Tool to search the ASV sequences against a reference database and available whole metagenomic sequencing (WMS) from a selected number of nasal washes (n = 6) to identify *Haemophilus* species. The Luminex xMAP

METHODS

Study populations

The SAGE and GALA II studies are 2 ongoing case-control studies approved by the Human Research Protection Program Institutional Review Board of the University of California, San Francisco (San Francisco, Calif).[E1] The ethics approval numbers are 210362 (SAGE) and 217802 (GALA II). Whereas GALA II recruited Hispanic and Latino individuals, SAGE II focused on African Americans. For the GALA II and SAGE studies, the participants (aged 8-21 years) declared that all 4 of their grandparents were of Hispanic/Latino or African American ancestry, respectively. All participants or parents provided written consent. The same sample collection procedures, protocols, and questionnaires were used for both studies. The SAGE participants were recruited in California's San Francisco Bay Area. The GALA II participants were recruited in different centers in the United States as well as in Puerto Rico. Asthma was diagnosed by the attending physician and on the basis of reports of symptoms or use of controller or rescue medication in the 2 years preceding enrollment. Asthma exacerbations were defined as the administration of oral corticosteroids, emergency asthma care, or hospitalizations in the 12 months preceding study enrollment.

Spirometry was performed with a KoKo PFT Spirometer (nSpire Health Inc, Louisville, Colo) according to American Thoracic Society recommendations (1995). Subjects with asthma were instructed to withhold their bronchodilator medications for at least 8 hours before pulmonary function testing. Baseline lung function values were normalized to obtain predicted values with the Global Lung Initiative 2012 reference equations.[E2] Measurements of IgE levels were conducted in duplicate by using the ImmunoCAP 100 system (Phadia, Kalamazooo, Mich).

Asthma severity was defined following the Expert Panel Report 3 guidelines for managing asthma based on the medication treatment.[E3] Classification of asthma control was based on the items included in the Childhood Asthma Control Test (C-ACT)[E4] and the Asthma Control Questionnaire[E5] based asthma nighttime symptoms, daytime symptoms, activity limitations, rescue medication use, and lung function measurements. The following questions were used to collect the data on each category: (1) In general during the past week, how much shortness of breath did the child experience because of asthma? (2) On average during the past week, how often was the child awakened by his or her asthma during the night? (3) In general during the past week, how limited was the child in his or her activities because of asthma? (4) How many times did the patient report short-acting β-agonist use in the past 2 weeks? (5) What is the FEV_1 percent predicted value and the ratio of FEV_1 to forced vital capacity? and (6) When was the child last prescribed oral steroids for asthma? The first 5 questions were scored as 0 (best asthma control), 1, or 2 (worst asthma control). Additionally, prescription of oral steroids for asthma was scored as 0 (>6 months ago) or 1 (≤6 months ago). The maximum score across the 5 domains was used to define asthma control with a score of 0 (very well controlled), 1 (not well controlled), or 2 (very poorly controlled).

Genotyping and statistical analysis

Genotyping was performed with the Axiom LAT1 array (World Array 4, Affymetrix, Santa Clara, Calif), as previously described,[E6,E7] and a subset of additional samples was also genotyped with the Axiom LAT1 Array Plus HLA.[E8] Quality control was performed as described elsewhere.[E6,E7] Imputation was carried out by using the Haplotype Reference Consortium[E9] as the reference population in the Michigan Imputation Server. Global ancestry proportions were estimated by using the software program ADMIXTURE,[E10] assuming African and European parental populations for African Americans and European, African, and Native American populations for Hispanics/Latinos. Reference haplotypes from parental populations were estimated from genotyping data from the European (CEU) and African (YRI) samples from the HapMap Project Phase III, which were downloaded from the Axiom Genotype Data Set (https://www.thermofisher.com/es/es/home/life-science/microarray-analysis/microarray-data-analysis/microarray-analysis-sample-data/axiom-genotype-data-set.html). For the Native American population, 71 individuals genotyped by using the Axiom LAT1 array (Affymetrix) were used as a reference, as described elsewhere.[E7] Association of genetic variants and phenotypes was tested through the logistic Wald test for binary traits and linear Wald test for quantitative measures by using EPACTS, version 3.2.6.[E11] In the discovery stage, the number of independent variants was determined by using PLINK,[E12] version 1.9, through a pruning procedure to exclude variants in linkage disequilibrium, retaining those with an r^2 value less than 0.2 within a 50-kb window and 5-SNP window shift for each candidate region. This r^2 threshold was selected on the basis of the findings of Sobota et al[E13] for populations of African ancestry indicating that other thresholds are overly conservative and result in a reduced statistical power to detect associations in these populations. Meta-analyses were performed with the METASOFT software tool.[E14] Genetic heterogeneity was assessed by means of the I^2 and Cochran Q tests. Power calculations were performed by using the software CATs, version 0.0.2,[E15] for a 2-stage case-control study, assuming a severe asthma exacerbation prevalence of 55%, as observed in the populations analyzed and different values of minor allele frequency (range 10%-50%) and genotype relative risk (range 1.2-2.0). *In silico* validation of the significant genetic variants in European populations was performed using public online databases, including data from the UK Biobank based on the SAIGE analysis of International Classification of Diseases (ICD)-derived traits[E16] and the Neale lab GWAS round 2[E17], as well as data from the Michigan Genomics Initiative.[E18]

In silico functional evaluation of the variant associated

Evidence of significant expression quantitative trait loci and *cis*-splicing quantitative trait loci was queried in the GTEx, version 8, database.[E19] Lung tissue and fibroblasts were considered biologically relevant tissues and cells for asthma exacerbations,[E20] given its involvement in airway inflammation and remodeling,[E21,E22] which has an important role in this phenotype.

REFERENCES

E1. Nishimura KK, Galanter JM, Roth LA, Oh SS, Thakur N, Nguyen EA, et al. Early-life air pollution and asthma risk in minority children. The GALA II and SAGE II studies. Am J Respir Crit Care Med 2013;188:309-18.

E2. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. Eur Respir J 2012;40:1324-43.

E3. National Asthma Education and Prevention Program. Expert Panel Report 3 (EPR-3): Guidelines for the Diagnosis and Management of Asthma-Summary Report 2007. J Allergy Clin Immunol 2007;120:S94-138.

E4. Liu AH, Zeiger R, Sorkness C, Mahr T, Ostrom N, Burgess S, et al. Development and cross-sectional validation of the Childhood Asthma Control Test. J Allergy Clin Immunol 2007;119:817-25.

E5. Juniper EF, O'Byrne PM, Guyatt GH, Ferrie PJ, King DR. Development and validation of a questionnaire to measure asthma control. Eur Respir J 1999;14:902-7.

E6. White MJ, Risse-Adams O, Goddard P, Contreras MG, Adams J, Hu D, et al. Novel genetic risk factors for asthma in African American children: precision medicine and the SAGE II study. Immunogenetics 2016;68:391-400.

E7. Pino-Yanes M, Gignoux CR, Galanter JM, Levin AM, Campbell CD, Eng C, et al. Genome-wide association study and admixture mapping reveal new loci associated with total IgE levels in Latinos. J Allergy Clin Immunol 2015;135:1502-10.

E8. Genes-environments and Admixture in Latino Asthmatics (GALA II) GWAS Study; 2015. Available at: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs001274.v1.p1. Accessed May 12, 2019.

E9. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet 2016;48:1279-83.

E10. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res 2009;19:1655-64.

E11. Kang HM. EPACTS (Efficient and Parallelizable Association Container Toolbox); 2016. Available at: https://genome.sph.umich.edu/wiki/EPACTS. Accessed August 10, 2019.

E12. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4.

E13. Sobota RS, Shriner D, Kodaman N, Goodloe R, Zheng W, Gao YT, et al. Addressing population-specific multiple testing burdens in genetic association studies. Ann Hum Genet 2015;792:136-47.

E14. Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. Am J Hum Genet 2011;88:586-98.

E15. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet 2006;38:209-13.

E16. UKBiobank ICD PheWeb. Analysis of 1403 ICD-based traits using SAIGE; 2020. Available at: http://pheweb.sph.umich.edu/SAIGE-UKB/pheno/495.2. Accessed June 11, 2020.

E17. GWAS round 2 results. Neale lab; 2018. Available at: http://www.nealelab.is/uk-biobank. Accessed June 11, 2020.

E18. MGI freeze2 PheWeb; 2020. Available at: http://pheweb.sph.umich.edu/pheno/495.2. Accessed June 11, 2020.

E19. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013;45:580-5.

E20. Westergren-Thorsson G, Larsen K, Nihlberg K, Andersson-Sjöland A, Hallgren O, Marko-Varga G, et al. Pathological airway remodelling in inflammation. Clin Respir J 2010;4S1:1-8.

E21. Descalzi D, Folli C, Scordamaglia F, Riccio AM, Gamalero C, Canonica GW. Importance of fibroblasts-myofibroblasts in asthma-induced airway remodeling. Recent Pat Inflamm Allergy Drug Discov 2007;1:237-41.

E22. Reeves SR, Barrow KA, Kolstad TK, White MP, Rich LM, Wight TN, et al. Fibroblast gene expression following asthmatic bronchial epithelial cell conditioning correlates with epithelial donor lung function and exacerbation history. Sci Rep 2018;8:15768.

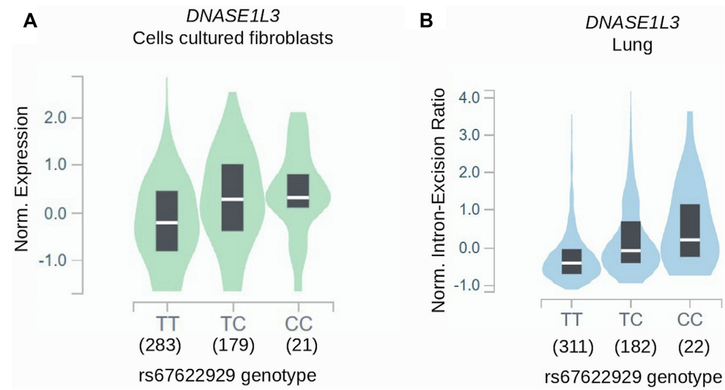


FIG E1. A, Expression quantitative trait loci (eQTL) analysis of SNP rs67622929 for *DNASE1L3* in cell-cultured fibroblasts from the GTEx,[E19] version 8, portal ($P = 7.1 \times 10^{-16}$). **B,** Splicing quantitative trait loci (sQTL) analysis of rs67622929 for *DNASE1L3* in lung from GTEx,[E19] version 8, portal. The splicing affects the intron 58204881:58208218:clu_47729 ($P = 9.4 \times 10^{-9}$).

TABLE E1. Estimation of the statistical power for the genomic regions assessed under different GRR scenarios taking into account the number of independent SNPs tested within each genomic region in the 2-stage case-control study

Minor allele frequency	Stage	No. of independent SNPs	GRR = 1.2	GRR = 1.4	GRR = 1.6	GRR = 1.8	GRR = 2.0
10%	Discovery	348	**91**	**100**	**100**	**100**	**100**
	Replication	1	47	**100**	**100**	**100**	**100**
	Joint analysis	1	47	**100**	**100**	**100**	**100**
20%	Discovery	348	**100**	**100**	**100**	**100**	**100**
	Replication	1	77	**100**	**100**	**100**	**100**
	Joint analysis	1	77	**100**	**100**	**100**	**100**
30%	Discovery	384	**100**	**100**	**100**	**100**	**100**
	Replication	1	**87**	**100**	**100**	**100**	**100**
	Joint analysis	1	**87**	**100**	**100**	**100**	**100**
40%	Discovery	348	**100**	**100**	**100**	**100**	**100**
	Replication	1	**89**	**100**	**100**	**100**	**100**
	Joint analysis	1	**89**	**100**	**100**	**100**	**100**
50%	Discovery	348	**100**	**100**	**100**	**100**	**100**
	Replication	1	**88**	**100**	**100**	**100**	**100**
	Joint analysis	1	**88**	**100**	**100**	**100**	**100**

GRR, Genotype relative risk.
Statistical power was obtained by using CATs, version 0.0.2, comparing different minor allele frequency values and GRRs, for a mean prevalence of severe asthma exacerbations of 55%, as observed in the populations analyzed in this case-control study. Values of statistical power of 80% or higher than are highlighted in boldface.

TABLE E2. Summary of association results for the 6 genes analyzed in SAGE

Gene	*ALPL*	*CLC*	*CPA3*	*CXCR2*	*DNASE1L3*	*IL1B*
SNPs tested (no.)	789	743	709	582	710	556
Independent SNPs (no.)	100	72	77	40	57	38
Bonferroni threshold	5.00×10^{-4}	6.94×10^{-4}	6.49×10^{-4}	1.25×10^{-3}	8.77×10^{-4}	1.32×10^{-3}
Significant SNPs (no.)	0	0	0	0	1	0
SNP with minimum *P* value	rs10799704	rs1568101	rs2001228	rs56979819	rs67622929	rs3136557
Minimum *P* value	5.08×10^{-4}	9.18×10^{-4}	1.43×10^{-4}	0.014	7.93×10^{-4}	8.8×10^{-3}

The Bonferroni-like threshold was calculated as the 0.05/number of independent variants located within each gene limit plus or minus 100 kb.

TABLE E3. Association results (additive model) for the C allele of rs67622929 with different outcomes

	SAGE		GALA II		Meta-analysis		
Trait	**OR/β (95% CI)***	***P* value**	**OR/β (95% CI)***	***P* value**	**OR/β (95% CI)***	***P* value**	**Cochran Q *P* value**
Asthma exacerbations (adjusted age, sex, and global ancestry)	1.48 (1.17-1.86)	7.9×10^{-4}	1.18 (1.01-1.37)	.03	1.26 (1.11-1.43)	2.6×10^{-4}	.103
FEV_1 % predicted	–1.17 (–1.90 to –0.45)	.104	–0.45 (–0.97 to 0.07)	.391	–0.70 (–1.12 to –0.27)	.099	.418
FVC % predicted	–0.93 (–1.62 to –0.24)	.178	0.05 (–0.48 to 0.58)	.930	0.05 (–0.33 to 0.42)	.900	1.000
FEV_1/FVC % predicted	–0.81 (0.17 to –1.79)	.107	–0.25 (–0.78 to 0.28)	.345	–0.38 (–0.84 to 0.86)	.110	.331
Total serum IgE levels	0.03 (–0.02 to 0.07)	.538	0.02 (0.00-0.04)	.491	0.02 (0.00-0.04)	.365	.855

FVC, Forced vital capacity.
*β-estimates or ORs are reported for quantitative measurement and binary traits, respectively.

TABLE E4. Association results between rs67622929 and susceptibility to asthma with exacerbations in public online databases including individuals of European descent

Study	Phenotype*	No. of cases	No. of controls	β (C allele)	SE	*P* value	Details
UK Biobank	Asthma with exacerbations vs in subjects without asthma	252	375,505	–0.05	0.11	.67	Based on data imputed to the HRC panel and SAIGE analysis of ICD-derived traits[E16]
UK Biobank	Asthma-related hospitalizations vs in subjects without asthma	1,986	359,208	-1.13×10^{-4}	2.12×10^{-4}	.59	Imputed genotypes from the HRC plus UK10K and KGP reference panels (Neale laboratory GWAS round 2)[E17]
Michigan Genomics Initiative	Asthma with exacerbations vs in subjects without asthma	1,506	32,282	–0.05	0.046	.20	Based on MGI data (freeze2) imputed to the HRC panel[E18]

GWAS, Genome-wide association study; *HRC*, Haplotype Reference Consortium; *ICD*, International Classification of Diseases; *KGP*, 1000 Genomes project; *MGI*, Michigan Genomics Initiative; *SE*, SE of the β-estimate.

*The time period for which the presence or absence of exacerbations is determined is not detailed in the databases from which the information was extracted.

TABLE E5. *cis*-Acting expression quantitative trait loci effects described at the GTEx, version 8, portal for rs67622929

Gene symbol	*P* value	NES	Tissue
DNASE1L3	7.10×10^{-16}	0.48	Cells, cultured fibroblasts
DNASE1L3	3.90×10^{-15}	0.65	Testis
DNASE1L3	1.20×10^{-11}	0.36	Adipose, subcutaneous
DNASE1L3	1.00×10^{-9}	0.35	Thyroid
ABHD6	6.70×10^{-8}	0.15	Skin, sun-exposed (lower leg)
PDHB	1.20×10^{-7}	0.34	Brain, cortex
RP11-456N14.4	3.30×10^{-7}	0.35	Heart, left ventricle
RP11-456N14.4	7.00×10^{-7}	0.33	Adipose, visceral (omentum)
DNASE1L3	9.50×10^{-7}	0.32	Skin, not sun-exposed (suprapubic)
ABHD6	1.20×10^{-6}	0.23	Testis
DNASE1L3	1.70×10^{-6}	0.27	Skin, sun-exposed (lower leg)
ABHD6	2.90×10^{-6}	0.21	Esophagus, muscularis
DNASE1L3	4.70×10^{-6}	0.28	Breast, mammary tissue
ABHD6	4.80×10^{-6}	0.14	Skin, not sun-exposed (suprapubic)
RPP14	1.70×10^{-5}	0.14	Muscle, skeletal
DNASE1L3	2.20×10^{-5}	0.29	Nerve, tibial
PDHB	3.30×10^{-5}	0.25	Brain, cerebellar hemisphere
PDHB	4.00×10^{-5}	0.08	Muscle, skeletal
FLNB-AS1	4.40×10^{-5}	0.22	Heart, atrial appendage
FLNB-AS1	8.70×10^{-5}	0.13	Skin, not sun-exposed (suprapubic)
DNASE1L3	1.00×10^{-4}	0.23	Heart, atrial appendage
FLNB-AS1	1.30×10^{-4}	0.34	Brain, cerebellum
FLNB-AS1	2.10×10^{-4}	0.14	Thyroid
DNASE1L3	2.30×10^{-4}	0.26	Esophagus, muscularis
RP11-456N14.4	2.50×10^{-4}	0.22	Adipose, subcutaneous

NES, Normalized effect size.
The NES magnitude has no direct biologic interpretation at the GTEx portal.[E19]

TABLE E6. Splicing quantitative trait effects described at the GTEx v8 Portal for rs67622929

Gene symbol	Intron ID	*P* value	NES	Tissue
DNASE1L3	58204881:58208218:clu_48244	7.60×10^{-34}	0.75	Thyroid
DNASE1L3	58204881:58208218:clu_38886	2.00×10^{-26}	1.1	Spleen
DNASE1L3	58204881:58208218:clu_42898	2.50×10^{-21}	0.95	Pituitary
DNASE1L3	58204881:58208218:clu_35842	1.50×10^{-16}	0.83	Adrenal gland
DNASE1L3	58204881:58208218:clu_30751	1.10×10^{-11}	0.52	Liver
DNASE1L3	58204881:58208218:clu_39805	2.90×10^{-9}	0.43	Heart, atrial appendage
DNASE1L3	58204881:58208218:clu_47729	9.40×10^{-9}	0.37	Lung
DNASE1L3	58204881:58208218:clu_45777	1.90×10^{-8}	0.44	Breast, mammary tissue
DNASE1L3	58204881:58208218:clu_41846	3.70×10^{-8}	0.35	Esophagus, mucosa

ID, Identifier; *NES*, normalized effect size.
The NES magnitude has no direct biologic interpretation in GTEx.[E19]

TABLE E7. Association results of rs67622929 (C allele) with asthma exacerbations stratified by ethnicity

Population	MAF case patients (%)	MAF controls (%)	OR (95% CI)	*P* value
African Americans in SAGE	21.16	15.61	1.48 (1.18-1.87)	7.93×10^{-4}
Puerto Ricans in GALA II	21.08	21.47	0.97 (0.77-1.23)	.823
Mexicans in GALA II	33.96	27.96	1.30 (1.00-1.67)	.043
Other Latinos in GALA II	30.30	27.26	1.35 (0.97-1.88)	.076

MAF, Minor allele frequency.

4.3 Genome-wide association study reveals a novel locus for asthma with severe exacerbations in diverse populations

The extensive lack of representation of ethnically diverse populations in biomedical research, and more specifically in genetic studies, prompted us to focus on analysing the genetic variation involved in AEs in minority populations at high risk of asthma. This third *Chapter* presents the results of GWAS of asthma with severe exacerbations in ethnically diverse children and youth using a specific phenotyping approach. In this analysis, controls were defined as non-asthmatic subjects, whereas cases comprised asthma patients with severe exacerbations.

A total of twelve variants were associated with asthma with severe exacerbations in Hispanics/Latinos at the suggestive significance level ($p \leq 5x10^{-5}$). One variant located nearby the long intergenic non-protein coding RNA 1913 (*LINC01913*) gene exhibited a significant and consistent effect in Hispanics/Latinos and African Americans. Additionally, this variant was associated with the protein kinase domain containing, cytoplasmic (*PKDCC*) gene DNA methylation in whole-blood from Hispanics/Latinos and *LINC01913* gene expression in lung tissues, according to publicly available data. This is one of the few studies evaluating the genetic factors involved in asthma with exacerbations in individuals from admixed populations.

This *Chapter* was published as an Original Article entitled 'Genome-wide association study reveals a novel locus for asthma with severe exacerbations in diverse populations', contributor Maria Pino-Yanes [Herrera-Luis E, Espuela-Ortiz A, Lorenzo-Diaz F, Keys KL, Mak ACY, Eng C, Huntsman S, Villar J, Rodriguez-Santana JR, Burchard EG, Pino-Yanes M] in *Pediatric Allergy and Immunology*, volume 32, and issue 1 in 2021 (doi: 10.1111/pai.13337). Copyright © 2021 John Wiley and Sons Ltd. and the European Academy of Allergy and Clinical Immunology. Published by John Wiley and Sons Ltd. This article is reproduced under the terms of John Wiley and Sons Ltd. and the European Academy of Allergy and Clinical Immunology license.

Received: 26 May 2020 | Revised: 8 July 2020 | Accepted: 11 August 2020
DOI: 10.1111/pai.13337

ORIGINAL ARTICLE

WILEY

Genome-wide association study reveals a novel locus for asthma with severe exacerbations in diverse populations

Esther Herrera-Luis[1] | Antonio Espuela-Ortiz[1] | Fabian Lorenzo-Diaz[1] | Kevin L. Keys[2,3] | Angel C. Y. Mak[2] | Celeste Eng[2] | Scott Huntsman[2] | Jesús Villar[4,5,6] | Jose R. Rodriguez-Santana[7] | Esteban G. Burchard[2,8] | Maria Pino-Yanes[1,4,9]

[1]Genomics and Health Group, Department of Biochemistry, Microbiology, Cell Biology and Genetics, Universidad de La Laguna, San Cristóbal de La Laguna, Tenerife, Spain

[2]Department of Medicine, University of California San Francisco, San Francisco, CA, USA

[3]Berkeley Institute for Data Science, University of California Berkeley, Berkeley, CA, USA

[4]CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain

[5]Multidisciplinary Organ Dysfunction Evaluation Research Network, Research Unit, Hospital Universitario Dr Negrín, Las Palmas de Gran Canaria, Spain

[6]Keenan Research Center for Biomedical Science at the Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, ON, Canada

[7]Centro de Neumología Pediátrica, San Juan, Puerto Rico

[8]Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, CA, USA

[9]Instituto de Tecnologías Biomédicas (ITB), Universidad de La Laguna, San Cristóbal de La Laguna, Santa Cruz de Tenerife, Spain

Correspondence
Maria Pino-Yanes, Genomics and Health Group, Department of Biochemistry, Microbiology, Cell Biology and Genetics, Universidad de La Laguna, Apartado 456, 38200, San Cristóbal de La Laguna, Santa Cruz de Tenerife, Spain.
Email: mdelpino@ull.edu.es

Funding information
This study was funded by the Spanish Ministry of of Science, Innovation, and Universities, the State Research Agency, and the European Regional Development Funds from the European Union (MICIU/AEI/FEDER, UE, SAF2017-83417R). EH-L was funded by a fellowship (PRE2018-083837) from the Spanish Ministry of Science, Innovation, and Universities. MP-Y was supported by the Ramón y Cajal Program from the Spanish Ministry of Economy, Industry, and Competitiveness (RYC-2015-17205). KLK was supported by the National Heart, Lung, and Blood Institute (NHLBI) (R01HL135156-S1), the UCSF Bakar Computational Health Sciences Institute, the Gordon and Betty Moore Foundation (GBMF3834), and the Alfred P. Sloan Foundation grant 2013-10-27 to UC Berkeley through the Moore-Sloan Data Sciences Environment initiative at the

Abstract

Background: Severe asthma exacerbations are a major cause of asthma morbidity and increased healthcare costs. Several studies have shown racial and ethnic differences in asthma exacerbation rates. We aimed to identify genetic variants associated with severe exacerbations in two high-risk populations for asthma.

Methods: A genome-wide association study of asthma in children and youth with severe exacerbations was performed in 1283 exacerbators and 2027 controls without asthma of Latino ancestry. Independent suggestive variants ($P \le 5 \times 10^{-6}$) were selected for replication in 448 African Americans exacerbators and 595 controls. Case-only analyses were performed comparing the exacerbators with additional 898 Latinos and 524 African Americans asthma patients without exacerbations, while adjusting by treatment category as a proxy of asthma severity. We analyzed the functionality of associated variants with in silico methods and by correlating genotypes with methylation levels in whole blood in a subset of 473 Latinos.

Results: We identified two genome-wide significant associations for susceptibility to asthma with severe exacerbations, including a novel locus located at chromosome 2p21 (rs4952375, odds ratio = 1.39, $P = 3.8 \times 10^{-8}$), which was also associated with asthma exacerbations in a case-only analysis (odds ratio = 1.25, $P = 1.95 \times 10^{-3}$). This polymorphism is an expression quantitative trait locus of the long intergenic non-protein coding RNA 1913 (*LINC01913*) in lung tissues ($P = 1.3 \times 10^{-7}$) and influences

© 2020 European Academy of Allergy and Clinical Immunology and John Wiley & Sons Ltd

Berkeley Institute for Data Science (BIDS). Financial support to conduct part of this work onsite at Universidad de La Laguna (ULL) was supported by ULL Campus America 2019 Initiative.
The Genes-Environments and Admixture in Latino Americans (GALA II) Study and the Study of African Americans, Asthma, Genes and Environments (SAGE) were supported by the National Heart, Lung, and Blood Institute (NHLBI) of the National Institutes of Health (NIH) (R01HL117004 and X01HL134589). Additional study enrollment was supported by the Sandler Family Foundation, the American Asthma Foundation, the Amos Medical Faculty Development Program from the Robert Wood Johnson Foundation, the Harry Wm. and Diana V. Hind Distinguished Professorship in Pharmaceutical Sciences II, and the National Institute of Environmental Health Sciences (NIEHS) (R01ES015794). The generation, cleaning, quality control, and analysis of GALA II and SAGE data was funded by the NHLBI (R01HL128439, R01HL135156, R01HL141992, and R01HL141845); the NIEHS (R21ES24844); the National Institute on Minority Health and Health Disparities (NIMHD) (P60MD006902, R01MD010443, and R56MD013312); the National Institute of General Medical Sciences (NIGMS) (RL5GM118984); the Tobacco-Related Disease Research Program (24RT-0025 and 27IR-0030); and the National Human Genome Research Institute (NHGRI) (U01HG009080) to EGB.

Editor: Jon Genuneit

methylation levels of the protein kinase domain-containing cytoplasmic (*PKDCC*) gene in whole-blood cells ($P = 9.8 \times 10^{-5}$).
Conclusion: We identified a novel susceptibility locus for severe asthma exacerbations in Hispanic/Latino and African American youths with functional effects in gene expression and methylation status of neighboring genes.

KEYWORDS
African American, asthma, ethnic differences, exacerbations, gene expression, genome-wide association study, Hispanic, Latino, methylation, single nucleotide polymorphism, susceptibility

1 | INTRODUCTION

Asthma is a chronic inflammatory disease influenced by both genetic and environmental factors. Despite treatment with controller medications, some patients experience severe asthma exacerbations, which are defined as episodes requiring emergency care, hospitalizations, or the use of systemic corticosteroids to prevent a serious or fatal outcome.[1] These can be triggered by environmental influences, such as viral infections, fungal allergens, or air pollution.[2] Genetic factors also predispose patients to higher susceptibility to asthma exacerbations.[2] In children, these exacerbations are associated with lower quality of life,[3] a progressive loss of lung function,[4] and an increase in their parents' work absence.[5] Therefore, asthma exacerbations represent a major component of the economic burden of asthma.

Among pediatric populations in the United States, asthma prevalence and asthma-related emergency department or urgent care center visits are highest in African American and Puerto Rican populations.[6] Moreover, African ancestry has been associated with a higher risk of developing asthma among African Americans[7] and Puerto Ricans,[8] as well as with an increased risk of exacerbations among African Americans.[9,10] These data point to a genetic component potentially contributing to the disparities in asthma exacerbations.

The genetic determinants of severe asthma exacerbations have been studied mostly in European populations, focusing on genes

> **Key Message**
>
> The first genomic analysis comparing asthma cases with severe exacerbations against controls performed in Hispanic/Latino and African American revealed a novel genome-wide significant susceptibility locus. The associatedvariant demonstrated to have functional consequences in methylation and gene expression in blood and lung tissue, respectively. Our results provide two different gene targets of interest for asthma exacerbations whose specific role should be further explored in future functional analysis.

previously associated with asthma susceptibility through candidate gene association studies.[11] The only four genome-wide association studies (GWAS) focused on asthma exacerbations performed to date suggest that loci for exacerbations differ from those of asthma susceptibility.[12-15] One GWAS analyzed extreme phenotypes by comparing exacerbators and children without asthma and revealed a novel locus at the cadherin-related family member 3 (*CDHR3*) gene, not associated in any GWAS focused on asthma susceptibility.[12] The other three GWAS of asthma exacerbations performed case-only analyses, comparing asthma patients with and without exacerbations. One study identified several single nucleotide polymorphisms (SNPs) in alpha-T-catenin (*CTNNA3*) significantly associated with asthma exacerbations at genome-wide significant level.[13] Two studies flagged *CMTR114* and *APOBEC3B-APOBEC3C*[15] as suggestively associated with asthma exacerbations despite treatment with inhaled corticosteroids.

We hypothesized that a GWAS comparing asthma patients with severe exacerbations with non-asthmatic controls could reveal novel specific association signals in two high-risk populations for asthma. We aimed to identify genetic variants associated with severe asthma exacerbations in Hispanics/Latinos and African Americans.

2 | METHODS

2.1 | Study populations

The Genes-Environments and Admixture in Latino Americans (GALA II) and the Study of African Americans, Asthma, Genes, and Environments (SAGE) are two case-control studies of pediatric asthma in minority populations of the United States.[16] Both studies were approved by the Human Research Protection Program Institutional Review Board of the University of California, San Francisco. Participants/parents provided written assent/consent, respectively. Asthma was defined by physician diagnosis, use of controller or rescue medication, and report of symptoms in the 2 years preceding enrollment. Controls were excluded if they reported asthma, rhinitis, hay fever, allergy, or eczema, wheezing, shortness of breath, or use of medication for allergies. Asthma exacerbations were defined as the presence of any of the following asthma-related events in the 12 months before enrollment: administration of oral corticosteroids (OCS), seeking emergency care, and hospitalizations.[1] A total of 1283 exacerbators and 2027 controls from GALA II were analyzed in the discovery stage. In the replication stage, 448 exacerbators and 595 controls from SAGE were included. Additionally, asthma patients without exacerbations (898 from GALA II and 524 from SAGE) were used as controls in case-only analyses.

2.2 | Genotyping and imputation

Genotyping was conducted with the Axiom LAT1 array and with the Axiom LAT1 Array Plus HLA (Affymetrix), and the SNPs and individuals resulting from the intersection of the two arrays were subjected to quality control. The Haplotype Reference Consortium (HRC)[17] r1.1 was used as imputation reference panel. Quality control and imputation are detailed in the Supporting Information.

2.3 | Statistical analysis

The association between 9 737 707 SNPs with minor allele frequency (MAF) ≥1% and imputation quality score $R^2 \geq 0.3$, and asthma with severe exacerbations was tested through logistic regression models using EPACTS 3.2.6,[18] with correction for age, sex, and the first two genotype principal components (PCs), which explained most genetic variation (Figure S1). Conditional analyses were performed with GCTA-COJO v1.92.2[19] using the default options to identify independent variants ($r^2 \leq 0.9$) within 10 Mb distance from the lead SNP at each locus with suggestive level of significance ($P \leq 5 \times 10^{-6}$). These variants were examined for replication in SAGE, and shared variants were meta-analyzed with METASOFT.[20] Random or fixed-effects models were chosen based on the heterogeneity estimated by Cochrane's Q test. Genome-wide significance was declared with $P \leq 5 \times 10^{-8}$ after the meta-analysis across both stages.

To refine the SNPs whose effect was driven by severe asthma exacerbations or asthma, we conducted a case-only analysis by comparing exacerbators and non-exacerbators with correction for age, sex, two PCs, and disease severity defined based on treatment category (Table 1). For the region associated with asthma exacerbations in this analysis, fine mapping was conducted within 1 Mb flanking the lead SNP in Latinos using three imputation panels: HRC, 1000 Genomes Project (1KGP) Phase III, and the Consortium on Asthma among African-Ancestry Populations in the Americas (CAAPA). For each SNP, imputed dosages were selected based on the highest R^2.

In secondary analyses, we evaluated the association of the most significant SNP with lung function measurements by regression models adjusted for the first two PCs and asthma status. Moreover, we performed stratified analyses to assess the potential confounding effect of atopic comorbidities and asthma control. See Supporting Information for more details.

2.4 | Methylation quantitative trait loci (cis-meQTL) analysis

We performed a *cis*-methylation quantitative trait locus (meQTL) analysis in 473 GALA II whole-blood samples (236 asthma cases and 237 controls) profiled with the Infinium HumanMethylation450 BeadChip array (Illumina, Inc). We tested for association with the M-values within 500 kb of the most strongly associated SNP with correction for sex, asthma, exacerbations, African and Native American ancestries, and inferred white cell counts. A false discovery rate (FDR) of 5% was used to declare significance accounting for the multiple comparisons performed. Detailed procedures are provided in the Supporting Information.

2.5 | In silico functional evaluation

Functional annotation was conducted with HaploReg v4.2[21] based on linkage disequilibrium (LD) data of admixed American populations from the 1KGP. Evidence of expression quantitative trait loci (eQTLs) was searched in the Genotype-Tissue Expression (GTEx)[22] v8.

3 | RESULTS

3.1 | Characteristics of study populations

The demographics and clinical characteristics of participants included in the analysis are shown in Table 1. Among asthma cases, Latinos sought unexpected asthma care more often than African

TABLE 1 Clinical and demographic characteristics of the subjects included in the case-control and case-only analyses

	GALA II			SAGE		
	Asthma with exacerbations (n = 1283)	Asthma without exacerbations (n = 898)	Control (n = 2027)	Asthma with exacerbations (n = 448)	Asthma without exacerbations (n = 524)	Control (n = 595)
Gender (% female)	582 (45.4)[a]	403 (44.9)	1137 (56.1)	214 (47.8)[a]	263 (50.2)	340 (57.1)
Mean age in years ± SD	12.2 ± 3.1[a,b]	13.4 ± 3.5	13.9 ± 3.6	13.3 ± 3.6[a,b]	14.4 ± 3.6	15.6 ± 3.8
African ancestry (%)	18.9 ± 13.2[a,b]	12.6 ± 11.9	14.2 ± 12.0[a]	83.6 ± 18.1	81.5 ± 21.5	81.7 ± 18.1
Native American ancestry (%)	22.9 ± 22.6[b]	37.4 ± 25.5	33.9 ± 27.8	NA	NA	NA
Asthma exacerbations in the last 12 mo, n (%)						
Acute asthma care	1092 (85.1)	NA	NA	231 (51.6)	NA	NA
Oral corticosteroid use	712 (55.5)	NA	NA	337 (75.2)	NA	NA
Hospitalizations	192 (15.0)	NA	NA	88 (19.6)	NA	NA
Asthma control, n (%)[c]						
Well controlled	210 (16.4)[b]	431 (48.0)	NA	60 (15.7)[b]	182 (36.8)	NA
Partially controlled	533 (41.5)[b]	282 (31.4)	NA	112 (29.4)	150 (30.4)	NA
Poorly controlled	540 (42.1)[b]	185 (20.6)	NA	209 (54.9)[b]	162 (32.8)	NA
Treatment category n (%)[d]						
Step 1	464 (36.4)[b]	505 (56.4)	NA	97 (21.7)[b]	242 (46.5)	NA
Step 2	417 (32.7)[b]	233 (26.1)	NA	182 (40.6)	210 (40.3)	NA
Step 3	394 (30.9)[b]	157 (17.5)	NA	169 (37.7)[b]	69 (13.2)	NA
Eczema, n (%)[e]	224 (17.7)	134 (15.1)	NA	188 (48.5)[b]	159 (32.4)	NA
Hay fever/rhinitis, n (%)[f]	564 (45.2)[b]	314 (35.4)	NA	148 (36.2)	142 (28.9)	NA
FEV_1 (%)[g]	97.9 ± 15.7[a,b]	103.6 ± 14.4	108.0 ± 12.6	91.2 ± 11.9	92.4 ± 11.6	96.1 ± 10.8
FVC (%)[g]	102.0 ± 16.0[a,b]	107.3 ± 15.3	108.0 ± 13.6	96.5 ± 11.8	97.0 ± 10.9	98.4 ± 11.8
FEV_1/FVC (%)[g]	96.4 ± 7.6[a]	97.2 ± 7.2	99.7 ± 6.7	94.1 ± 8.3	95.3 ± 7.6	97.8 ± 7.4

Note: For continuous variables, the mean and standard deviation are displayed, and the Mann-Whitney-Wilcoxon test was applied for the comparison of cases vs controls. For categorical variables, the number and proportion of subjects in each category are shown and a chi-square test was applied for the comparison of cases vs controls.

Abbreviations: FEV_1, forced expiratory volume in the first second; FVC, forced vital capacity; NA, not available/applicable; SD, standard deviation.

[a] $P < .01$ for the comparison between asthma cases with exacerbations and controls.

[b] $P < .01$ for the comparison between asthma cases with exacerbations and asthma cases without exacerbations.

[c] Asthma control was available for 2181 and 875 asthma cases from GALA II and SAGE, respectively.

[d] The severity medication regime was categorized into three levels based on the use of short beta-agonists (step 1), one inhaled corticosteroid, leukotriene inhibitor, or theophylline tablet (step 2), more than one or a combination of an inhaled corticosteroid, leukotriene inhibitor, or theophylline tablet or a combination of inhaled corticosteroids and long-beta-agonists (step 3). Medication data were available for 2170 and 969 asthma cases from GALA II and SAGE, respectively.

[e] Eczema status was available for 2155 and 879 subjects from GALA II and SAGE, respectively.

[f] Hay fever status was available for 2133 and 901 subjects from GALA II and SAGE, respectively.

[g] FEV_1, FVC, and FEV_1/FVC were available for 2698, 2710, and 2662 GALA II subjects, respectively, and 1058, 1075, and 1056 SAGE subjects, respectively.

Americans, while hospitalizations and OCS use were more frequent in African Americans.

3.2 | Discovery in Latinos from GALA II

The quantile-quantile plot showed slight genomic inflation (Figure S2A, λ_{GC} = 1.02), which was driven by the strong association signals in LD at the known asthma locus 17q12-q21 (Figure S2B). A total of 171 SNPs were suggestively associated with severe exacerbations ($P \le 5 \times 10^{-6}$) (Figure 1; Table S1). From these, conditional analyses identified 12 independent SNPs with imputation quality R^2 values ranging from 0.91 to 1.00 (Table 2), of which rs12946510 at the *GRB7/IKZF3* region showed the strongest association (odds ratio [OR] for T allele = 0.68, 95% confidence interval [CI]:0.60-0.76, $P = 6.82 \times 10^{-11}$).

3.3 | Replication in African Americans from SAGE and meta-analysis

Of the 12 SNPs selected for replication, 3 SNPs were unsuitable for replication in African Americans due to their low frequency in this population (MAF <1%). From the remaining SNPs (Table 2), rs4952375 at the 2p21 region near *LINC01913* replicated (OR for A allele = 1.53, 95% CI: 1.12-2.08, $P = 7.43 \times 10^{-3}$). A combined meta-analysis resulted in two genome-wide significant associations for rs12946510 (OR for C allele = 0.71, 95% CI: 0.64-0.79, $P = 1.12 \times 10^{-10}$) and rs4952375 (OR for A allele = 1.39, 95% CI: 1.23-1.56, $P = 3.79 \times 10^{-8}$) (Figure 2). We next assessed these two genome-wide significant associations in a case-only meta-analysis (n = 3139) including exacerbators and non-exacerbators (Table S2). The SNP rs4952375 was associated with increased risk of severe exacerbations among asthma patients (OR for A allele = 1.25, 95% CI: 1.09-1.44, $P = 1.95 \times 10^{-3}$). However, rs12946510 showed no association in the case-only analysis, suggesting that its effect is likely driven by asthma rather than severe exacerbations. The results of case-only analyses of the suggestive associations identified in the discovery are also provided in Table S2.

3.4 | Fine mapping of 2p21 and assessment of additional phenotypes

Since rs4952375 was associated with asthma exacerbations in the case-control and the case-only analyses and has not been identified by previous GWAS, we performed fine mapping with additional imputation reference panels. Despite increasing the number of variants analyzed from 4817 to 6906 (27.8% and 2.4% exclusive of 1KGP and CAAPA, respectively), no additional genome-wide or suggestively associated SNPs were found (Figure S3).

Further meta-analyses for asthma-related traits were conducted for the SNP rs4952375 (Table S3). We first evaluated whether the association found for asthma with exacerbations differed by asthma control status, stratifying the exacerbators as well controlled or poorly controlled and comparing them with the controls. Although the association was only significant among poorly controlled exacerbators, no heterogeneity was found for the effect of rs4952375 among the two groups when compared to the controls (Cochran's Q $P = 0.263$ and $P = 0.988$ for GALA II and SAGE, respectively). Additionally, given that almost half of the asthma patients reported eczema or rhinitis, we assessed whether atopic comorbidities could be a confounder performing stratified analyses (Table S3). No heterogeneity for the effect of rs4952375 was

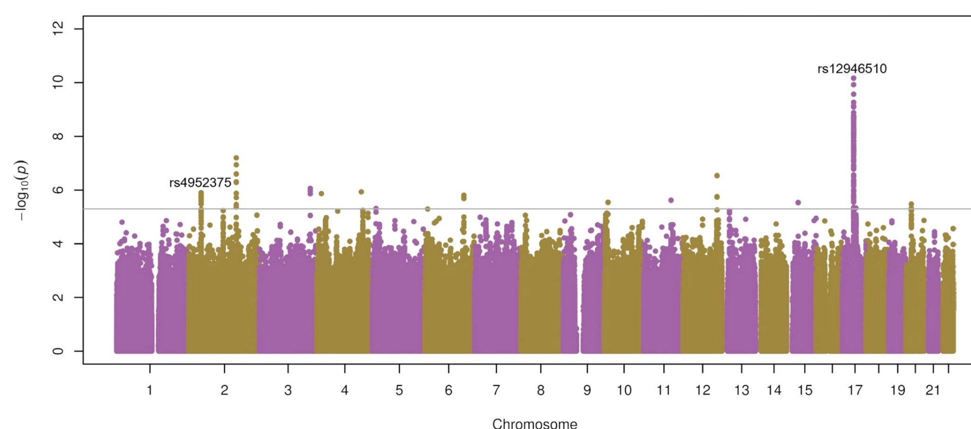


FIGURE 1 Manhattan plot of the GWAS for GALA II (represented as -log10 *P*-value on the *y*-axis) along the chromosomes (*x*-axis). The suggestive significance threshold for replication is indicated by the gray line ($P = 5 \times 10^{-6}$). GALA II, Genes-Environments and Admixture in Latino Americans; GWAS, genome-wide association studies [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Association and meta-analysis results for the independent SNPs identified in Hispanics/Latinos followed up for replication in African Americans

				GALA II (n = 3310)			SAGE (n = 1043)			Meta-analysis (n = 4353)		
rsID	Chr.: position (bp)[a]	A1/A2	Closest gene	Freq.[b]	OR (95% CI)[b]	*P*-value[b]	Freq.[b]	OR (95% CI)[b]	*P*-value[b]	OR (95% CI)[b]	*P*-value[b]	Model
rs4952375	2:42 016 863	A/T	*snoZ247/LINC01913*	0.23	1.37 (1.20-1.55)	1.24×10^{-6}	0.10	1.53 (1.12-2.08)	7.43×10^{-3}	1.39 (1.23-1.56)	3.79×10^{-8}	FE
rs17759529	2:162 851 013	G/C	*DPP4*	0.16	0.67 (0.58-0.77)	6.35×10^{-8}	0.14	0.91 (0.69-1.18)	.467	0.76 (0.57-1.02)	3.80×10^{-7}	RE2
rs187089444	3:174 519 230	A/C	*NAALADL2*	0.03	2.19 (1.60-2.99)	8.70×10^{-7}	NA	NA	NA	NA	NA	NA
rs181042383	4:14 887 486	A/G	*LINC00504/CPEB2-DT*	0.04	1.95 (1.49-2.56)	1.35×10^{-6}	NA	NA	NA	NA	NA	NA
rs78968521	4:152 214 673	G/A	*PRSS48*	0.02	2.38 (1.68-3.38)	1.16×10^{-6}	0.01	0.69 (0.31-1.54)	.360	1.35 (0.40-4.55)	1.45×10^{-5}	RE2
rs112599943	6:133 499 724	C/T	*EYA4*	0.04	1.92 (1.47-2.50)	1.55×10^{-6}	0.08	1.40 (1.00-1.95)	.047	1.70 (1.38-2.09)	6.00×10^{-7}	FE
rs76605301	10:12 227 306	T/C	*NUDT5*	0.03	2.14 (1.56-2.94)	2.87×10^{-6}	NA	NA	NA	NA	NA	NA
rs591709	11:93 520 315	T/C	*MED17*	0.15	0.70 (0.60-0.81)	2.39×10^{-6}	0.20	1.28 (1.02-1.60)	.037	0.94 (0.52-1.70)	9.65×10^{-6}	RE2
rs11613333	12:116 880 378	C/T	*MIR4472-2*	0.42	1.34 (1.20-1.50)	2.90×10^{-7}	0.08	1.22 (0.87-1.71)	.243	1.33 (1.20-1.48)	1.64×10^{-7}	FE
rs12946510	17:37 912 377	T/C	*GRB7/IKZF3*	0.32	0.68 (0.60-0.76)	6.82×10^{-11}	0.21	0.85 (0.67-1.07)	.164	0.71 (0.64-0.79)	1.12×10^{-10}	FE
rs113516984	17:45 964 442	C/T	*RP11-6N17.4/SP2*	0.11	0.66 (0.56-0.79)	4.64×10^{-6}	0.07	0.67 (0.43-1.02)	.064	0.66 (0.56-0.78)	7.82×10^{-7}	FE
rs6132057	20:18 390 980	G/A	*DZANK1*	0.15	0.69 (0.59-0.81)	3.33×10^{-6}	0.07	1.33 (0.93-1.91)	.117	0.94 (0.49-1.79)	2.58×10^{-5}	RE2

Abbreviations: A1, effect allele; A2, non-effect allele; Chr., chromosome; CI, confidence interval; FE, fixed-effects model; Freq., effect allele frequency; NA, not available/applicable; OR, odds ratio; RE2, Han and Eskin's random-effects models; rsID, reference SNP identifier.

[a]Coordinates are referred to the GRCh37 reference genome.

[b]Referred to the effect alleles (additive model).

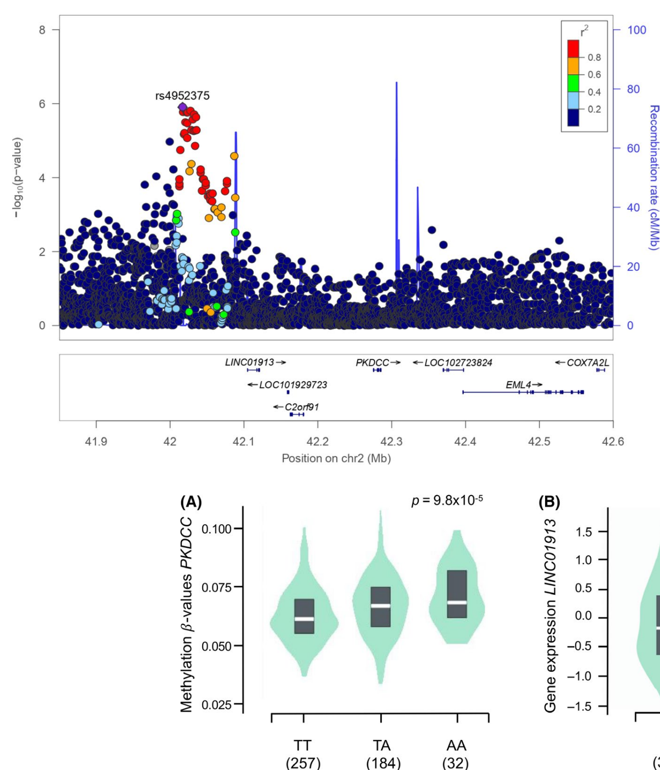


FIGURE 2 Regional plot of association in the discovery phase for the novel association detected at 2p21. The statistical significance of association results (-log10 *P*-value) is represented for each SNP as a dot (left *y*-axis) and recombination rate (right *y*-axis) by chromosome position (*x*-axis). SNPs are colored to show their LD with the top hit based on the pairwise r^2 values from the American admixed populations of the 1KGP. LD, linkage disequilibrium; SNP, single nucleotide polymorphism [Colour figure can be viewed at wileyonlinelibrary.com]

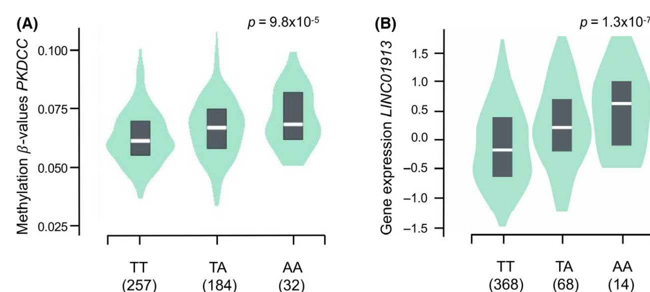


FIGURE 3 Functional analysis of the rs4952375. A, Boxplot of blood methylation levels at cg17892159 annotated to *PKDCC* by rs4952375 genotype in a subset of GALA II subjects (n = 473). B, Gene expression levels of *LINC01913* in lung tissue by genotype at rs4952375 indicating that this SNP is an eQTL, based on data obtained from GTEx (n = 450). Abbreviations: eQTL, expression quantitative trait loci; GALA II, Genes-Environments and Admixture in Latino Americans; SNP, single nucleotide polymorphism [Colour figure can be viewed at wileyonlinelibrary.com]

found between exacerbators with these morbidities vs controls or exacerbators without them when compared to controls (Cochran's Q P = 0.842 for both studies). Moreover, rs4952375 was not associated with predicted baseline forced expiratory volume in one second (FEV_1), forced vital capacity (FVC), or the FEV_1/FVC ratio, in an analysis corrected for asthma status and the top two PCs (Table S3).

3.5 | Functional analysis of the associated variant

In a subset of GALA II participants with methylation data available, we identified rs4952375 as a *cis*-meQTL for the CpG site cg17892159, which lies near the protein kinase domain-containing cytoplasmic (*PKDCC*) gene. Carriers of the risk allele showed increased methylation levels for cg17892159 ($P = 9.79 \times 10^{-5}$; FDR = 1.3%) (Figure 3A).

Additionally, rs4952375 is a lung eQTL for the uncharacterized long intergenic non-coding RNA (lncRNA) 1913 (*LINC01913*) in the GTEx database ($P = 1.3 \times 10^{-7}$). Lung *LINC01913* expression increases when the risk allele is present (Figure 3B). Furthermore, this variant is in high LD ($r^2 \geq 0.8$) with 31 lung eQTLs for *LINC01913* (Table S4).

3.6 | Validation of associations identified by previous GWAS

We assessed previous associations reported for asthma exacerbations[11] and moderate-to-severe asthma[23] for validation in GALA II and SAGE. We replicated the association of *GSDMB* and *IKZF3* at genome-wide significant level and 10 additional SNPs at nominal level with asthma with severe exacerbations (Table S5). However, only the SNP rs1837253 from *TSLP* was also nominally associated with severe exacerbations in a case-only analysis (Table S6).

4 | DISCUSSION

To our knowledge, this is the first GWAS of asthma with severe exacerbations in Latino and African American children and youth. We performed a two-stage GWAS and identified a novel locus for susceptibility to asthma with exacerbations that reached genome-wide significance in the combined meta-analysis. The strongest association at chromosome 2p21 is located at rs4952375, which was also associated with severe asthma exacerbations in a case-only meta-analysis of Latinos and African Americans. The sentinel SNP rs4952375 and its variants in LD ($r^2 \geq 0.8$) are located nearby *LINC01913*. Specifically, rs4952375 is located 87.8 kb away from the 5'UTR of *LINC01913*. In fact, the SNP rs4952375 and their proxies in LD regulate gene expression of the long non-coding RNA *LINC01913* in the lung. We found that the A allele of rs4952375 is associated with a higher risk of asthma exacerbations and increased gene expression of *LINC01913* in the lung. Long non-protein coding RNAs have been previously implicated in asthma phenotypes and other lung diseases as they interact with microRNAs and transcription factors involved in different biologic processes.[24] Our results suggest that lncRNAs could also contribute to the pathogenesis of asthma exacerbations, although the mechanism involving *LINC01913* in asthma remains unknown.

Our findings indicate that rs4952375 is also a meQTL for a CpG site annotated to *PKDCC*, a gene located 258 kb away from the associated marker. Genetic variation from *PKDCC* or its neighbor gene encoding for the echinoderm microtubule-associated protein-like 4 (*EML4*) has been associated with lung function measurements[25] and atopy.[26] However, those polymorphisms were not associated with severe asthma exacerbations in our study and are not in LD with rs4952375 (Table S7), suggesting that they are independent association signals. Likewise, rs4952375 was not associated with lung function measurements in our study and did not show differences between severe exacerbators with or without atopic comorbidities.

PKDCC is a tyrosine-protein kinase that regulates cell adhesion through signal transduction,[27] including phosphorylation of *MMP1*,[28] which is involved in airway remodeling and whose activity is increased during exacerbations in relation to the exacerbation severity.[29] Based on the biologic processes ontology described in the database Open Targets,[30] *PKDCC* is involved in cell differentiation, lung alveolus development, multicellular organism growth, and bone mineralization. Furthermore, mice deficient in *Pkdcc* show lung abnormalities.[30] Our results suggest that the risk allele for asthma exacerbations is associated with higher methylation levels of *PKDCC*, which would imply that its gene expression could be downregulated in individuals with asthma exacerbations.

Hispanics/Latinos are descendants of the admixture of European, African, and Native American populations, with large variability in the ancestral proportions depending on the historical events that each particular subgroup has undergone.[31] On the other hand, African Americans derive from the admixture of European and African populations, with large proportions of African ancestry, with mean values ranging from 73% to 93%.[31] The high asthma prevalence among African-admixed individuals, particularly in developed countries,[6-10] suggests a shared genetic basis that may be a risk factor for asthma under environmental influences that are absent in non-Westernized societies.[8] In this work, we leveraged this specific genetic background to identify a novel locus for severe asthma exacerbations. Given the different ancestral composition, the risk allele of rs4952375 was more frequent in Latinos than in African Americans (22.7% vs 10.1%), in agreement with 1KGP (Figure S4). However, the association had consistent effects in both populations.

We also attempted to validate previous SNPs associated with asthma exacerbations or moderate-to-severe asthma in Latinos and African Americans. Although we found replication of several associations, no evidence of association was detected with severe asthma exacerbations in a case-only analysis, except for one SNP, suggesting that the effect of those variants is due to asthma susceptibility rather than exacerbations. Therefore, our findings suggest that the genetic determinants of severe exacerbations may differ from those involved in moderate-to-severe asthma. Alternatively, the lack of replication could be due to the differences in the characteristics of the patients analyzed in the current study compared to Shrine et al,[23] in terms of ancestry, age, and study design. Of note, the SNPs rs6967330, located in the human rhinovirus C receptor *CDHR3* gene and previously associated with asthma with severe exacerbations,[12] did not replicate in our study, which could reflect different pathogenic mechanisms between the two studies, which differed in age, ethnicity, and definition of exacerbations.

We acknowledge some limitations in this study. First, reporting of asthma exacerbations was retrospective in both studies, which could introduce some imprecision into the phenotypes analyzed. Second, the fact that several SNPs are in strong LD in the region complicates the identification of the causal variant(s). Third, asthma is a complex disease resulting from the interaction between genetic and environmental exposures, which could be critical for the development of exacerbations, and those exposures were not addressed in our study. Fourth, limited functional data were available within the samples analyzed. However, data available on public databases (GTEx[22] and OpenTargets[30]) provided evidence about the functional role of the associated variant regulating gene expression and the importance of the *PKDCC* in lung development based on knockout mice models.

On the other hand, this study has several strengths. Foremost, this is the first GWAS of asthma with severe exacerbations specifically focused on minority racial/ethnic groups at high risk of asthma morbidity and underrepresented in biomedical studies. Second, we showed that examining a specific asthma phenotype revealed genetic variants that have not been identified by previous GWAS of asthma susceptibility on the same populations. Third, we performed an in-depth fine mapping of the region of interest using a combination of three different panels and assessed the association with several clinical subtypes.

In summary, we conducted a GWAS of asthma with severe exacerbations in Hispanic/Latino and African American children and youth. Our study identified a novel genome-wide significant association for severe asthma exacerbations on chromosome 2p21. The associated variant is a lung eQTL for the *LINC01913* gene and is also a whole-blood meQTL of a CpG site annotated to *PKDCC*. Additional functional follow-up is required to confirm the role of this locus in asthma exacerbations.

ACKNOWLEDGEMENTS

The GALA II and SAGE study collaborators include Shannon Thyne, UCSF; Harold J. Farber, Texas Children's Hospital; Denise Serebrisky, Jacobi Medical Center; Rajesh Kumar, Lurie Children's Hospital of Chicago; Emerita Brigino-Buenaventura, Kaiser Permanente; Michael A. LeNoir, Bay Area Pediatrics; Kelley Meade, UCSF Benioff Children's Hospital, Oakland; William Rodriguez-Cintron, VA Hospital, Puerto Rico; Pedro C. Avila, Northwestern University; Jose R. Rodríguez-Santana, Centro de Neumología Pediátrica; Luisa N. Borrell, City University of New York; Adam Davis, UCSF Benioff Children's Hospital, Oakland; Saunak Sen, University of Tennessee; and Fred Lurmann, Sonoma Technologies, Inc; Michael A. LeNoir, Bay Area Pediatrics; Kelley Meade, UCSF Benioff Children's Hospital, Oakland.

The authors would like to thank the families and patients for their participation and thank the numerous healthcare providers and community clinics for their support and participation. In particular, the authors thank Sandra Salazar, the study coordinator, and the recruiters who obtained the data: Duanny Alva, Gaby Ayala-Rodriguez, Lisa Caine, Elizabeth Castellanos, Jaime Colon, Denise DeJesus, Blanca Lopez, Brenda Lopez, Louis Martos, Vivian Medina, Juana Olivo, Mario Peralta, Esther Pomares, Jihan Quraishi, Johanna Rodriguez, Shahdad Saeedi, Dean Soto, and Ana Taveras.

CONFLICTS OF INTEREST

EH-L reports a fellowship from the Spanish Ministry of Science, Innovation, and Universities. KLK reports grants from National Heart, Lung, and Blood Institute (NHLBI), the Gordon and Betty Moore Foundation to UC Berkeley, the UCSF Bakar Computational Health Sciences Institute, and the Alfred P. Sloan Foundation to UC Berkeley. KLK also reports non-financial support from Universidad de La Laguna Campus America 2019 Initiative. MP-Y and FLD report grants from Spanish Ministry of Science, Innovation, and Universities, the State Research Agency, and the European Regional Development Funds from the European Union (MICIU/AEI/FEDER, UE) and MP-Y reports grants from the Spanish Ministry of Economy, Industry, and Competitiveness. EGB reports grants from the National Institutes of Health, the Tobacco-Related Disease Research Program, the Sandler Family Foundation, the American Asthma Foundation, the Amos Medical Faculty Development Program from the Robert Wood Johnson Foundation and from the Harry Wm. and Diana V. Hind Distinguished Professorship in Pharmaceutical Sciences II. The rest of authors declare no conflict of interest.

AUTHOR CONTRIBUTION

Esther Herrera-Luis: Conceptualization (supporting); Data curation (equal); Formal analysis (lead); Methodology (equal); Writing-original draft (lead); Writing-review & editing (equal). **Antonio Espuela-Ortiz:** Formal analysis (equal); Methodology (supporting); Writing-original draft (supporting); Writing-review & editing (supporting). **Fabian Lorenzo-Diaz:** Conceptualization (supporting); Funding acquisition (supporting); Writing-original draft (supporting); Writing-review & editing (supporting). **Kevin L. Keys:** Data curation (supporting); Formal analysis (supporting); Writing-review & editing (supporting). **Angel CY Mak:** Data curation (supporting); Formal analysis (supporting); Writing-review & editing (supporting). **Celeste Eng:** Data curation (supporting); Investigation (supporting); Writing-review & editing (supporting). **Scott Huntsman:** Data curation (lead); Methodology (supporting); Writing-review & editing (supporting). **Jesús Villar:** Conceptualization (supporting); Writing-review & editing (supporting). **José R Rodriguez-Santana:** Investigation (supporting); Writing-review & editing (supporting). **Esteban G. Burchard:** Conceptualization (supporting); Funding acquisition (equal); Supervision (supporting); Writing-review & editing (supporting). **Maria Pino-Yanes:** Conceptualization (lead); Funding acquisition (equal); Methodology (equal); Supervision (lead); Writing-original draft (supporting); Writing-review & editing (lead).

ORCID

Esther Herrera-Luis https://orcid.org/0000-0003-4150-5454

Antonio Espuela-Ortiz https://orcid.org/0000-0003-0572-9549

Fabian Lorenzo-Diaz https://orcid.org/0000-0002-3398-198X
Kevin L. Keys http://orcid.org/0000-0003-4886-6666
Angel C. Y. Mak http://orcid.org/0000-0002-5372-4198

Jesús Villar https://orcid.org/0000-0001-5687-3562
Esteban G. Burchard http://orcid.org/0000-0001-7475-2035

Maria Pino-Yanes https://orcid.org/0000-0003-0332-437X

REFERENCES

1. Reddel HK, Taylor DR, Bateman ED, et al. Asthma control and exacerbations - standardizing endpoints for clinical asthma trials and clinical practice. *Am J Respir Crit Care Med*. 2009;180:59-99.
2. Cook J, Saglani S. Pathogenesis and prevention strategies of severe asthma exacerbations in children. *Curr Opin Pulm Med*. 2016;22:25-31.
3. Chipps BE, Haselkorn T, Rosén K, Mink DR, Trzaskoma BL, Luskin AT. Asthma exacerbations and triggers in children in TENOR: impact on quality of life. *J Allergy Clin Immunol Pract*. 2018;6:169-176.e2.
4. Calhoun WJ, Haselkorn T, Miller DP, Omachi TA. Asthma exacerbations and lung function in patients with severe or difficult-to-treat asthma. *J Allergy Clin Immunol*. 2015;136:1125-1127.e4.
5. Sullivan PW, Ghushchyan V, Navaratnam P, et al. The national burden of poorly controlled asthma, school absence and parental work loss among school-aged children in the United States. *J Asthma*. 2018;55:659-667.
6. Oraka E, Iqbal S, Flanders WD, Brinker K, Garbe P. Racial and ethnic disparities in current asthma and emergency department visits: findings from the national health interview survey, 2001–2010. *J Asthma*. 2013;50:488-496.

7. Flores C, Ma S-F, Pino-Yanes M, et al. African ancestry is associated with asthma risk in African Americans. *PLoS One*. 2012;7:e26807.
8. Pino-Yanes M, Thakur N, Gignoux CR, et al. Genetic ancestry influences asthma susceptibility and lung function among Latinos. *J Allergy Clin Immunol*. 2015;135:228-235.
9. Rumpel JA, Ahmedani BK, Peterson EL, et al. Genetic ancestry and its association with asthma exacerbations among African American subjects with asthma. *J Allergy Clin Immunol*. 2012;130:1302-1306.
10. Grossman NL, Ortega VE, King TS, et al. Exacerbation-prone asthma in the context of race and ancestry in Asthma Clinical Research Network trials. *J Allergy Clin Immunol*. 2019;144:1524-1533.
11. Herrera-Luis E, Hernandez-Pacheco N, Vijverberg SJ, Flores C, Pino-Yanes M. Role of genomics in asthma exacerbations. *Curr Opin Pulm Med*. 2019;25:101-112.
12. Bønnelykke K, Sleiman P, Nielsen K, et al. A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. *Nat Genet*. 2014;46:51-55.
13. McGeachie MJ, Wu AC, Tse SM, et al. CTNNA3 and SEMA3D: promising loci for asthma exacerbation identified through multiple genome-wide association studies. *J Allergy Clin Immunol*. 2015;136:1503-1510.
14. Dahlin A, Denny J, Roden DM, et al. CMTR1 is associated with increased asthma exacerbations in patients taking inhaled corticosteroids. *Immun Inflamm Dis*. 2015;3:350-359.
15. Hernandez-Pacheco N, Farzan N, Francis B, et al. Genome-wide association study of inhaled corticosteroid response in admixed children with asthma. *Clin Exp Allergy*. 2019;49:789-798.
16. Nishimura KK, Galanter JM, Roth LA, et al. Early-life air pollution and asthma risk in minority children. The GALA II and SAGE II studies. *Am J Respir Crit Care Med*. 2013;188:309-318.
17. McCarthy S, Das S, Kretzschmar W, et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*. 2016;48:1279-1283.
18. Kang HM.EPACTS (Efficient and Parallelizable Association Container Toolbox). Date last updated: February 19, 2019. https://genome.sph.umich.edu/wiki/EPACTS. Accessed June 4, 2019.
19. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88:76-82.
20. Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am J Hum Genet*. 2011;88:586-598.
21. Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res*. 2016;44:D877-D881.
22. Lonsdale J, Thomas J, Salvatore M, et al. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*. 2013;45:580-585.
23. Shrine N, Portelli MA, John C, et al. Moderate-to-severe asthma in individuals of European ancestry: a genome-wide association study. *Lancet Respir Med*. 2019;7:20-34.
24. Wang Z, Ji N, Chen Z, et al. Next generation sequencing for long non-coding RNAs profile for CD4+ T cells in the mouse model of acute asthma. *Front Genet*. 2019;10:545.
25. Kichaev G, Bhatia G, Loh P-R, et al. Leveraging polygenic functional enrichment to improve GWAS power. *Am J Hum Genet*. 2019;104:65-75.
26. Castro-Giner F, Bustamante M, Ramon González J, et al. A pooling-based genome-wide analysis identifies new potential candidate genes for atopy in the European Community Respiratory Health Survey (ECRHS). *BMC Med Genet*. 2009;10:128.
27. Maddala R, Skiba NP, Rao PV. Vertebrate lonesome kinase regulated extracellular matrix protein phosphorylation, cell shape, and adhesion in trabecular meshwork cells. *J Cell Physiol*. 2017;232:2447-2460.
28. Bordoli M, Yum J, Breitkopf S, et al. A secreted tyrosine kinase acts in the extracellular environment. *Cell*. 2014;158:1033-1044.
29. Naveed S-u-N, Clements D, Jackson DJ, et al. Matrix metalloproteinase-1 activation contributes to airway smooth muscle growth and asthma severity. *Am J Respir Crit Care Med*. 2017;195:1000-1009.
30. Koscielny G, An P, Carvalho-Silva D, et al. Open Targets: a platform for therapeutic target identification and validation. *Nucleic Acids Res*. 2017;45:D985-D994.
31. Bryc K, Durand EY, Macpherson JM, Reich D, Mountain JL. The genetic ancestry of African Americans, Latinos, and European Americans across the United States. *Am J Hum Genet*. 2015;96:37-53.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Herrera-Luis E, Espuela-Ortiz A, Lorenzo-Diaz F, et al. Genome-wide association study reveals a novel locus for asthma with severe exacerbations in diverse populations. *Pediatr Allergy Immunol*. 2021;32:106–115. https://doi.org/10.1111/pai.13337

Link to Supporting Information:

https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2F pai.13337&file=pai13337-sup-0001-Supinfo.doc

4.4 Admixture mapping of severe asthma exacerbations loci in children and youth

This *Chapter* includes the results from the first admixture mapping of severe AEs in Hispanic/Latino children and youth ever conducted. Leveraging local ancestry, we revealed several significant genomic regions where local ancestry was associated with severe AEs in Mexican Americans and Puerto Ricans. Fine mapping of common and rare genetic variation within the admixture mapping peak revealed genetic variants that showed significant and consistent effects in both Latino/Hispanic ethnic subgroups. Additionally, genetic variants were followed up for trans-ancestral replication in Singaporean Chinese, African Americans, and European-descent populations. The effect on gene expression and DNA methylation of genetic variants that exhibited significant and consistent effects among Latino ethnic subgroups was investigated in whole-blood from Latino patients with asthma.

A total of six and three genomic regions showed a significant association of local ancestry with severe AEs in Mexican Americans and Puerto Ricans, respectively. In Mexican Americans, fine-mapping of common variants revealed two SNPs that showed significant and consistent effects in Puerto Ricans, and exhibited functional effects on expression of secretoglobin Family 3A Member 2 (*SCGB3A2*) gene and DNA methylation nearby the Dihydropyrimidinase Like 3 (*DPYSL3*) gene in whole-blood. None of these two common variants showed significant association in non-Hispanic/Latino populations, suggesting that these variants may exert ethnic-specific genetic effects.

This *Chapter* has been submitted as an Original Article entitled 'Admixture mapping of severe asthma exacerbations in Latino children and youth' to a peer-reviewed journal.

Admixture mapping of severe asthma exacerbations in Latino children and youth

Running title: Admixture mapping of asthma exacerbations in children

Esther Herrera-Luis, MSc[1], Angel C. Y. Mak, PhD[2], Javier Perez-Garcia, PharmD[1], Elena Martin-Gonzalez, MSc[1], Celeste Eng, BS[2], Kenneth B. Beckman, PhD[3], Scott Huntsman, MSc[2], Donglei Hu, PhD[2], Ruperto González-Pérez, MD, PhD[4,5], José M. Hernández-Pérez, MD, PhD[6,7], Elena Mederos-Luís, MD[4], Yang Yie Sio, PhD[8], Paloma Poza-Guedes, MD, PhD[4,5], Olaia Sardón, MD, PhD[9,10], Paula Corcuera, MD, PhD[9], Inmaculada Sánchez-Machín, MD, PhD[4], Javier Korta-Murua, MD, PhD[9,10], Carlos Martínez-Rivera, MD, PhD[11,12], Joaquim Mullol, MD, PhD[11,13], Xavier Muñoz, MD, PhD[11,14], Antonio Valero, MD, PhD[11,15], Joaquín Sastre, MD, PhD[11,16], Judith Garcia-Aymerich, PhD[17,18,19], Sabrina Llop, PhD[17,20], Maties Torrent, PhD[21], Maribel Casas, PhD[18,19], José R. Rodríguez-Santana, MD, FAAP, FCCP[22], Jesús Villar, MD, PhD, FCCM[11,23], Victoria del Pozo, PhD[11,24], Fabian Lorenzo-Diaz, PhD[1,25], L. Keoki Williams, MD, MPH[26], Erik Melén, MD, PhD[27,28], Fook Tim Chew, PhD[8], Luisa N. Borrell, DDS, PhD[29], Esteban G. Burchard, MD, MPH[3,30†], Maria Pino-Yanes, PhD[1,11,31†*]

From

[1]Genomics and Health Group, Department of Biochemistry, Microbiology, Cell Biology and Genetics, Universidad de La Laguna (ULL), La Laguna, Tenerife, Spain;

[2]Department of Medicine, University of California San Francisco, San Francisco, California, U.S.A.;

[3]UMN Genomics Center, Minneapolis, Minnesota, U.S.A.;

[4]Allergy Department, Hospital Universitario de Canarias, Santa Cruz de Tenerife, Tenerife, Spain;

[5]Asthma Unit, Hospital Universitario de Canarias, La Laguna, Tenerife, Spain;

[6]Pulmonary Medicine, Hospital Universitario de N.S de Candelaria, Santa Cruz de Tenerife, Spain;

[7]Pulmonary Medicine, Hospital General de La Palma, La Palma, Santa Cruz de Tenerife, Spain;

[8]Department of Biological Sciences, National University of Singapore, Singapore;

[9]Division of Pediatric Respiratory Medicine, Hospital Universitario Donostia, San Sebastián, Spain;

[10]Department of Pediatrics, University of the Basque Country (UPV/EHU), San Sebastián, Spain;

[11]CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain;

[12]Servicio de Neumología, Hospital Universitario Germans Trias i Pujol, Universitat Autònoma de Barcelona, Badalona, Spain;

[13]Rhinology Unit & Smell Clinic, ENT Department; Clinical & Experimental Respiratory Immunoallergy (IDIBAPS), Universitat de Barcelona, Barcelona, Spain;

[14]Servicio de Neumología, Hospital Vall d'Hebron, Barcelona, Spain;

[15]Allergy Unit & Severe Asthma Unit, Pneumonology and Allergy Department, Hospital Clínic; IDIBAPS; Universitat de Barcelona.Barcelona, Spain;

[16]Allergy Department, Hospital Universitario Fundación Jiménez Díaz, Madrid, Spain;

[17]Spanish Consortium for Research on Epidemiology and Public Health (CIBERESP), Madrid, Spain;

[18]ISGlobal, Barcelona, Spain;

[19]Universitat Pompeu Fabra, Barcelona, Spain;

[20]Epidemiology and Environmental Health Joint Research Unit, FISABIO−Universitat Jaume I−Universitat de València, Valencia, Spain;

[21]Ib-salut, Area de Salut de Menorca, Menorca, Spain;

[22]Centro de Neumología Pediátrica, San Juan, Puerto Rico;

[23]*Multidisciplinary Organ Dysfunction Evaluation Research Network, Research Unit, Hospital Universitario Dr. Negrín, Las Palmas de Gran Canaria, Spain;*

[24]*Immunology Department, Instituto de Investigación Sanitaria Hospital Universitario Fundación Jiménez Díaz, Madrid, Spain;*

[25]*Instituto Universitario de Enfermedades Tropicales y Salud Pública de Canarias (IUETSPC), Universidad de La Laguna (ULL), La Laguna, Tenerife, Spain;*

[26]*Center for Individualized and Genomic Medicine Research, Department of Internal Medicine, Henry Ford Health System, Detroit, MI, U.S.A.;*

[27]*Department of Clinical Sciences and Education, Södersjukhuset, Karolinska Institutet, Stockholm, Sweden;*

[28]*Sachs' Children's Hospital, South General Hospital, Stockholm, Sweden;*

[29]*Department of Epidemiology & Biostatistics, Graduate School of Public Health & Health Policy, City University of New York, New York, NY, U.S.A.;*

[30]*Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, California, U.S.A.;*

[31]*Instituto de Tecnologías Biomédicas (ITB), Universidad de La Laguna (ULL), La Laguna, Tenerife, Spain.*

[†] *Equal contribution*

***Corresponding author:**

Dr. María Pino-Yanes
Genomics and Health Group. Department of Biochemistry, Microbiology, Cell Biology and Genetics. Universidad de La Laguna (ULL)
Apartado 456, 38200, La Laguna, Tenerife, Spain
Phone: (+34) 922 316502 6343
Fax: (+34) 922 318 490
e-mail: mdelpino@ull.edu.es

Abstract count: 250

Word count: 3,507

Figures: 3

Tables: 4

Supplementary figures: 5

Supplementary tables: 10

FUNDING

This work was funded by the Spanish Ministry of Science and Innovation MCIN/AEI/10.13039/501100011033 and the European Regional Development Fund “ERDF A way of making Europe” by the European Union grant [SAF2017-83417R], and by MCIN/AEI/10.13039/501100011033 [grant PID2020-116274RB-I00]. Whole-genome sequencing for the TOPMed (Trans-Omics in Precision Medicine) program was supported by the NHLBI. Whole-genome sequencing for GALA II (NHLBI TOPMed: Genes-environments and Admixture in Latino Americans) Study (phs000920) was performed at the New York Genome Center [3R01HL117004-01S3]. Centralized read mapping and genotype calling, along with variant quality metrics and filtering, were provided by the TOPMed Informatics Research Center [3R01HL-117626-02S1]. Phenotype harmonization, data management, sample-identity quality control, and general study coordination were provided by the TOPMed Data Coordinating Center [3R01HL-120393-02S1]. WGS of part of GALA II was performed by New York Genome Center under The Centers for Common Disease Genomics of the Genome Sequencing Program (GSP) Grant [UM1 HG008901]. The GSP Coordinating Center [U24 HG008956] contributed to cross-program scientific initiatives and provided logistical and general study coordination. GSP is funded by the National Human Genome Research Institute, the National Heart, Lung, and Blood Institute, and the National Eye Institute [NA]. This work was also supported by the Sandler Family Foundation[NA], the American Asthma Foundation [NA], the RWJF Amos Medical Faculty Development Program [NA], Harry Wm. and Diana V. Hind Distinguished Professor in Pharmaceutical Sciences II [NA], the National Heart, Lung, and Blood Institute of the National Institutes of Health [R01HL117004, R01HL128439, R01HL135156, X01HL134589, R01HL155024, R01AI079139, R01HL141845, R01HL118267 and R01DK113003], the National Institute of Health and Environmental Health Sciences [R01ES015794 and R21ES24844], the National Institute on Minority Health and Health Disparities [R01MD010443 and R56MD013312], the Tobacco-Related Disease Research Program [Award Number 24RT-0025 and 27IR-0030]. The BAMSE study was funded by the Swedish Heart-Lung Foundation and the Swedish Research Council and Region Stockholm (ALF project and database maintenance). The INMA study was funded by the Instituto de Salud Carlos III [Red INMA G03/176, CB06/02/0041], Spanish Ministry of Health [FIS-FEDER PI16/1288, FIS-FEDER PI19/1338; Miguel Servet FEDER 15/0025 and 20/0006], Generalitat Valenciana [BEST/2020/059]. This work was also funded by Fundación Canaria Instituto de Investigación Sanitaria de Canarias [PIFIISC19/17]. JV was funded by ISCIII and the European Regional Development Fund “ERDF A way of making Europe” by the European Union [PI16/00049 and PI19/00141]. MP-Y was funded by the Ramón y Cajal Program [RYC-2015-17205] by MCIN/AEI/10.13039/501100011033 and by the European Social Fund “ESF Investing in your future” and by the EAACI Allergopharma Award 2021. JV and M.P.-Y were funded Instituto de Salud Carlos III (ISCIII) [CB06/06/1088]. EH-L was supported by a fellowship awarded by MCIN/AEI/10.13039/501100011033 and by “ESF Investing in your future” [PRE2018-083837]. JP-G was supported by a fellowship awarded by the Spanish Ministry of Universities [FPU19/02175].

CONFLICTS OF INTEREST

EH-L, and M.P.-Y. report funding from the Spanish Ministry of Science and Innovation (MCIN/AEI/10.13039/501100011033) and by “ESF Investing in your future” by the European Union. JP-G reports funding from the Spanish Ministry of Universities. M.P.-Y. and F.L.D. report grants from MCIN/AEI/10.13039/501100011033 and the European Regional Development Fund “ERDF A way of making Europe” by the European Union. MP-Y reports grant support from GlaxoSmithKline, Spain paid to Fundación Canaria Instituto de Investigación Sanitaria de Canarias (FIISC) for a project outside the submitted work. MP-Y and JV reports grants from Instituto de Salud Carlos III,

Madrid, Spain. JV also reports funding by ISCIII and the European Regional Development Fund "ERDF A way of making Europe". JMH-P has received fees from CSL Behring, GSK, Astra-Zeneca, laboratorios Menarini, Boehringer Ingelheim, FAES, laboratorios Esteve, Laboratorios Ferrer, Mundipharma, Laboratorios Rovi, Roche, Novartis, GRIFOLS, Pfizer, Acthelion-Jansen, Chiesi y Laboratorios Bial for the realization of courses, talks, consultancies, and other activities related to his professional activity. FTC has received research support from the Singapore Ministry of Education Academic Research Fund, Singapore Immunology Network (SIgN), National Medical Research Council (NMRC) (Singapore), Biomedical Research Council (BMRC) (Singapore), and the Agency for Science Technology and Research (A*STAR) (Singapore). F.T.C. has received consulting fees from Sime Darby Technology Centre; First Resources Ltd; Genting Plantation, and Olam International, outside the submitted work. YYS has received research support from the NUS Resilience & Growth Postdoctoral Fellowships. The other authors declare no conflict of interest.

AUTHOR CONTRIBUTION

EH-L: Conceptualization (supporting), Data curation (equal), Formal analysis (lead), Methodology (equal), Writing original draft (lead), Writing review and editing (equal); ACYM: Data curation (equal), Formal analysis (supporting), Writing review and editing (supporting); JP-G: Investigation (supporting), Formal analysis (supporting), Writing review and editing (equal); EM-G: Investigation (supporting), Writing review and editing (supporting); CE: Investigation (supporting), Writing-review and editing (supporting); KBB: Investigation (supporting), Writing-review and editing (supporting); SH: Data curation (equal), Methodology (supporting)), Writing-review and editing (supporting); DH: Data curation (equal), Writing-review and editing (supporting); RG-P: Resources (equal), Investigation (supporting), Writing-review and editing (supporting); JMH-P: Resources (equal), Investigation (supporting), Writing-review and editing (supporting); EM-L: Resources (equal), Investigation (supporting), Writing-review and editing (supporting); YYS: Resources (equal), Investigation (supporting), Data curation (equal), Formal analysis (equal), Writing-review and editing (supporting); PP-G: Resources (equal), Investigation (supporting), Writing-review and editing (supporting); OS: Resources (equal),Investigation (supporting), Writing-review and editing (supporting); PC: Resources (equal), Investigation (supporting), Writing-review and editing (supporting); IS-M: Resources (equal), Investigation (supporting), Writing-review and editing (supporting); JK-M: Resources (equal), Investigation (supporting), Writing-review and editing (supporting); CMR: Investigation (supporting), Writing review and editing (supporting); JM: Investigation (supporting), Writing review and editing (supporting); XM: Investigation (supporting), Writing review and editing (supporting); AV: Investigation (supporting), Writing review and editing (supporting); JS: Investigation (supporting), Writing review and editing (supporting); JG-A: Investigation (supporting); Formal analysis (supporting), Writing review and editing (supporting); SL: Investigation (supporting); Formal analysis (supporting), Writing review and editing (supporting); MT: Investigation (supporting), Formal analysis (supporting); MC: Investigation (supporting), Formal analysis (supporting), Writing review and editing (supporting); JRR-S: Investigation (supporting), Writing-review and editing (supporting); JV: Conceptualization (supporting), Resources (supporting), Writing-review & editing (equal); VP: Resources (supporting), Investigation (supporting), Writing-review and editing (supporting); FL-D: Resources (equal), Conceptualization (supporting), Funding acquisition (supporting), Writing-review and editing (equal); Funding acquisition (lead), Writing-review & editing (supporting); LKW: Resources (equal), Funding acquisition (lead), Writing-review & editing (supporting); EM: Resources (equal), Writing-review & editing (supporting); FTC: Resources (equal), Writing-review & editing (supporting); LNB: Funding acquisition (equal), Writing-review & editing (supporting); EGB: Resources (equal), Conceptualization (lead), Funding acquisition (lead), Supervision (supporting), Writing-review & editing (supporting); MP-Y: Resources (equal), Conceptualization (lead), Supervision (lead), Funding acquisition (lead); Methodology (equal), Writing original draft (supporting), Writing review and editing (lead).

ABSTRACT

Background: In the United States, genetically admixed populations have the highest asthma prevalence and severe asthma exacerbations rates. This could be explained by environmental factors, but also by ethnic-specific genetic variants. However, no admixture mapping has been performed for severe asthma exacerbations.

Objective: We sought to identify genetic variants associated with severe asthma exacerbations in Hispanics/Latinos by means of admixture mapping analyses and fine-mapping, and to assess their transferability to other populations and potential functional roles.

Methods: We performed an admixture mapping in 1,124 Puerto Rican and 625 Mexican American children with asthma. Fine-mapping of the significant peaks was performed via allelic testing of common and rare variants. We performed replication across Hispanic/Latino subgroups, and the transferability to non-Latino populations was assessed in 1,001 African Americans, 1,250 Singaporeans, and 941 Europeans with asthma. The effects of the variants on gene expression and DNA methylation from whole-blood were also evaluated.

Results: Genome-wide significant associations of Indigenous American ancestry with severe asthma exacerbations were found at 5q32 in Mexican Americans and at 13q13-q13.2 and 3p13 in Puerto Ricans. In Mexican Americans, a genome-wide significant association with European ancestry was also uncovered at 5q33.1 and 9q22.31. The SNPs rs28836559 (*STK32A*) and rs1144986 (*C5orf46*) showed consistent effects for severe asthma exacerbations across Hispanic/Latino subgroups, but were not validated in non-Latinos. These genetic variants regulated *DPYSL3* DNA methylation and *SCGB3A2* gene expression levels.

Conclusions: The first admixture mapping study of asthma exacerbations revealed novel regulatory genetic variants for *DPYSL3* and *SCGB3A2* that exhibited Latino-specific effects.

KEY MESSAGE

- **What is already known on this topic –** Admixed minorities in the United States, including Hispanics/Latinos, show disproportionate rates of asthma and asthma exacerbations compared to European-descent populations, but no study has assessed the contribution of genetic variation to asthma exacerbations by admixture mapping.
- **What this study adds –** The first admixture mapping of severe asthma exacerbations in Hispanics/Latinos revealed two ethnic-specific functional genetic variants with effects over two biologically plausibly genes implicated in this trait (*DPYSL3* and *SCGB3A2*).
- **How this study might affect research, practice or policy –** Our findings have prioritized two novel gene targets for future research in asthma.

KEYWORDS

Admixture mapping, Hispanics/Latinos, asthma, exacerbations, ancestry

ABBREVIATIONS

1KGP: 1000 Genomes project
95% CI: 95% confidence interval
CEU: Utah residents with Northern and Western European ancestry
eQTL: Expression quantitative trait loci
EPR-3: Expert Panel Report 3
FDR: False Discovery Rate
GALA II: Genes-Environment and Admixture in Latino Americans II study
GEMAS: Genomics and Metagenomics of Asthma Severity
GTEx: Genotype-Tissue Expression
GWAS: Genome-wide association study
Kb: Kilobase
MAF: Minor allele frequency
Mb: Megabase
MCMC: Markov Chain Monte Carlo
MEGA: MEchanism underlying the Genesis and evolution of Asthma
meQTL: Methylation quantitative trait loci
NAM: Native American
OR: Odds ratio
QC: Quality control
SAGE: Study of African Americans, Asthma, Genes and Environment
SCSGES: Singapore Cross Sectional Genetic Epidemiology Study
SNP: single nucleotide polymorphism
TOPMed: Trans-Omics for Precision Medicine
US: United States of America
WGS: Whole-genome sequencing
YRI: Yoruba in Ibadan, Nigeria

INTRODUCTION

Asthma is a chronic inflammatory disease that shows remarkable heterogeneity by ethnicity, and results from the complex interplay between environmental, behavioural, and genetic factors.[1] In fact, genetic ancestry is a critical factor for asthma susceptibility.[2] In the United States (U.S.), asthma prevalence is highest in Puerto Ricans (14.0%), followed by African Americans (10.7%), and lowest in Mexican Americans (5.4%) and Asian Americans (4.5%).[3] However, ethnic disparities are not limited to asthma prevalence, since Puerto Ricans and African Americans also show the highest rates of asthma-related emergency room or urgent care visits.[4] Asthma exacerbations, defined as those events requiring urgent asthma care, hospitalizations, or administration of systemic corticosteroids, contribute to the large healthcare expenditures of asthma[5] and asthma-related deaths.[6] Moreover, exacerbations impair the patients' quality of life[7] and long-term lung function.[8] Although environmental exposures, such as viral infections or air pollution, are known triggers for asthma exacerbations, several genetic loci for asthma exacerbations have been successfully uncovered through genome-wide associations studies (GWAS) in Hispanics/Latinos.[9–11]

Hispanics/Latinos are American admixed individuals with variable influences from Indigenous American, African and European populations (i.e., African ancestry is higher in Puerto Ricans compared to Mexicans, and Native American ancestry is higher in Mexicans).[12] In a scenario where a trait shows differential rates in admixed individuals, admixture mapping in the most affected populations can identify genetic variants associated with that trait. In fact, mapping causal variants by long-range linkage disequilibrium decreases multiple testing burden and can be coupled with functional approaches to prioritize genetic variants that could have population-specific effects.

Given the disparities in asthma prevalence and outcomes, multiple studies have leveraged locus-specific genetic ancestry to identify genetic variants associated with asthma-related outcomes.[12–15] However, no admixture mapping of severe asthma exacerbations has been performed to date. We hypothesized that genetic variation might partially explain the ethnic differences in severe asthma exacerbations among Hispanics/Latinos. Hence, we aimed to identify novel genetic variants associated with severe asthma exacerbations by admixture mapping in Hispanic/Latino children and youth with asthma. We then sought to evaluate the potential functional consequences of those genetic variants and attempted to validate them in non-Hispanic/Latino populations.

METHODS

A full description of the methods can be found in the Supplemental material.

Discovery population

We analysed genotypes from Puerto Ricans (n=1,124) and Mexican Americans (n=625) from the study of Genes-Environment and Admixture in Latino Americans (GALA II), a case-control study of asthma in Hispanics/Latinos aged 8-21 years-old recruited in the US and Puerto Rico between 2006 and 2014. GALA II was approved by the Human Research Protection Program Institutional Review Board of the University of California, San Francisco (US), and participants/parents provided written assent/consent, respectively. Participants were eligible if they reported four Hispanic/Latino grandparents. Asthma cases were diagnosed by a physician and had asthma symptoms and reported use of controller or rescue medication in the two years preceding enrolment.[16] Severe asthma exacerbations were defined as the presence of any of the following events in the past 12 months: use of oral corticosteroids, asthma-related hospitalizations, or unscheduled asthma care.[17] Treatment category, defined following the Expert Panel Report 3 (EPR-3) guidelines for the diagnosis and management of asthma,[18] was used as a proxy of asthma severity. Three levels were defined based on the use of short beta-agonists (step 1), one inhaled corticosteroid, leukotriene inhibitor, or theophylline tablet (step 2), more than one or a combination of an inhaled corticosteroid, leukotriene inhibitor, or theophylline tablet or a combination of inhaled corticosteroids and long-beta agonists (step 3).

Whole-genome sequencing data generation and processing

Whole-genome sequencing (WGS) was conducted on a HiSeq X system (Illumina, San Diego, CA). Data generation and quality control (QC) are described in the Supplemental material. Genotypes used in this study were based on TOPMed freeze 8 data.

Local and global ancestry estimation

The ancestral reference panel for ancestry estimation included 90 HapMap phase II Europeans (CEU), 90 Africans (YRI), and 71 Native Americans (NAM).[19] Local ancestry was estimated with RFMIX v2[20] based on biallelic single nucleotide polymorphisms (SNPs) with a PASS filter from the WGS that overlapped with the reference panel (402,838 SNPs). Global ancestry was calculated as the genome-wide average of local ancestry.

Admixture mapping in Hispanics/Latinos

The association of asthma exacerbations and the number of copies of African, European, and Indigenous American local ancestry (0, 1, or 2 copies) at each SNP analysed was tested separately in Latino/Hispanic subgroups. This was performed through logistic regression models with correction for global ancestry, age, sex, and treatment category (as a proxy for asthma severity). We estimated the effective number of tests using an empirical autocorrelation, to account for multiple comparison testing, as described in the Supplemental material.

Fine-mapping and replication

Fine-mapping of variants with minor allele frequency (MAF)≥1% within the genome-wide significant admixture mapping peaks was performed with correction for the same covariates used in the admixture mapping. To correct for multiple comparison, an adjusted threshold of significance was defined as p =<0.05/number of effective tests. To evaluate whether independent SNPs account for the admixture mapping signal, we performed step-wise conditional regression analyses adjusting the association of local ancestry by the allele dosage of associated SNPs. Interaction of local ancestry on the association of the SNPs with exacerbations was evaluated by a regression model including the same covariates that the main model and an interaction term SNP x local ancestry.

Peak-wise independent significant SNPs identified in Mexican Americans were replicated in Puerto Ricans and those identified in Puerto Ricans were assessed in Mexican Americans. Additional validation was sought in non-Latino participants with asthma, including 1,101 African Americans from the Study of African Americans, Asthma, Genes, and Environments (SAGE), 1,250 Singaporean Chinese from the Singapore Cross Sectional Genetic Epidemiology Study (SCSGES), and 941 Europeans from several studies: the MEchanism underlying the Genesis and evolution of Asthma (MEGA) study, the Genomics and Metagenomics of Asthma Severity (GEMAS) study, the Children, Allergy, Milieu, Stockholm, Epidemiology (BAMSE) study, and the Infancia and Medio Ambiente (INMA) study. Details of studies included in the replication are described in the Supplemental material and **Table E1**.

Fine-mapping of rare variants (MAF<1%) was conducted using the SKAT-O test[21] analyzing 1-kilobase (kb) sliding windows with 500-bp increments within the admixture mapping peak limits. The threshold of significance was defined based on Storey $q<0.05$.

Methylation profiling and quality control

DNA methylation from whole blood was profiled using the Illumina Infinium HumanMethylation450 BeadChip or the Infinium EPIC BeadChip arrays. QC is detailed in the Supplemental material. Briefly, low-quality probes and samples, outliers of DNA methylation, and samples with sex discordance or mixed genotype distributions on the control SNP probes were excluded. Standard background correction, dye-bias correction, inter-array normalization, and probe-type bias adjustment were conducted, and beta values were transformed to M-values.

RNA sample processing, sequencing, and quality control

Total RNA was quantified using the Quant-iT™ RiboGreen® RNA Assay Kit and normalized for library preparation with the Illumina TruSeq™ Stranded mRNA Sample Preparation Kit. Libraries were sequenced according to the manufacturer's protocols using the HiSeq 4000 system (Illumina). Sample processing and quality control are detailed in the Supplemental material. Expression values were normalized across samples using an inverse normal transformation.

Functional assessment of associated SNPs

Quantitative trait loci (QTL) analyses in data from whole blood were conducted separately in Mexican Americans and Puerto Ricans from GALA II for those CpG sites or genes with a transcription start site located within 1 Mb of the significant SNPs. Linear regression models were adjusted for exacerbations status, age, sex, treatment category, and genetic ancestry. *Cis-e*xpression QTL (eQTL) analyses were adjusted by the top 60 PEER factors[22], and *cis*-methylation QTL (meQTL) models by cell-type heterogeneity and methylation batch. The population sub-group results were then meta-analyzed. *In silico* evidence of functional effects of the variants on gene expression and DNA methylation was assessed using the Genotype-Tissue Expression (GTEx) v8 Portal,[23] PhenoScanner v2,[24] and the Genetics of DNA Methylation Consortium (GoDMC).[25] Chromatin interactions were determined using CHiCP.[26] Gene expression was inspected using the GTEx Portal[23] and REALGAR.[27]

RESULTS

Discovery sample characteristics

A total of 820 out of 1,124 Puerto Ricans had exacerbations in the last 12 months (72.9%), while in Mexican Americans 223 out of 625 individuals reported exacerbations for the same period (35.7%) Most of the exacerbators among Puerto Ricans and non-exacerbators among both ethnic subgroups were in the lowest treatment category, while the distribution in Mexican exacerbators was more balanced. Moreover, no significant differences in global ancestry were detected between exacerbators and non-exacerbators ($p>0.05$) (**Table 1**).

Stratified admixture mapping

A genome-wide significant association was found between Indigenous American ancestry and severe asthma exacerbations at chromosomes 3p13 and 13q13.2 in Puerto Ricans (**Table 2, Figure 1A, Figure E1** in the Supplemental material) ($p\leq1.16\text{x}10^{-4}$, accounting for 432 ancestry blocks). The most significant association at chromosome 3p13 corresponded to rs4677148, where Indigenous American ancestry was associated with lower odds of exacerbations (odds ratio [OR]: 0.57, 95% confidence interval [CI]: 0.44–0.74, $p=2.55\text{x}10^{-5}$), while European ancestry was associated with a higher risk of having exacerbations (OR: 1.40, 95% CI: 01.16–1.94, $p = 1.82\text{x}10^{-3}$) and no significant association with African ancestry was detected ($p>0.05$). The strongest association with the risk of severe exacerbations at 13q13.2 was located at rs10514839 (OR: 2.18, 95% CI: 1.55–3.06, $p=6.56 \text{ x } 10^{-6}$), whereby European or African ancestry showed no association.

Among Mexican Americans, we identified three admixture mapping peaks at chromosome 5q32 whereby Indigenous American ancestry was genome-wide significantly associated with the risk of severe exacerbations ($p\leq1.16\text{x}10^{-4}$, accounting for 431 ancestry blocks) (**Table 2, Figure 1B, Figure E2**). The most significant association in the region was located at rs10477350, where

Indigenous American was associated with higher odds of exacerbations (OR: 1.73, 95% CI: 1.34–2.24, p=3.23x10^{-5}) and European ancestry with lower risk (OR: 0.60, 95% CI: 0.46–0.78, p=1.46x10^{-4}). Moreover, European ancestry was genome-wide significantly associated (p≤1.35x10^{-4}, accounting for 470 ancestry blocks) with asthma exacerbations in two regions (**Table 2, Figure 1C**). The minimum *p* within the 5q33.1 peak was located at rs884083, being European ancestry associated with lower risk of having exacerbations (OR: 0.59, 95% CI: 0.45–0.77, p=1.25x10^{-4}) and Indigenous American ancestry with increased risk (OR: 1.60, 95% CI: 1.24–2.08, p = 3.48x10^{-4}). European ancestry at 9q22.2-q22.31 was also genome-wide significantly associated with lower odds of exacerbations (**Figure E3**), being the most significant association located at rs7034207 (OR: 0.58, 95% CI: 0.44–0.75, p=5.65x10^{-5} and OR: 1.50, 95% CI: 1.16–1.94, p=2.14x10^{-3}, for European and Indigenous American ancestry, respectively).

No genome-wide significant associations were found between severe asthma exacerbations and African ancestry in both ethnic subgroups or European ancestry in Puerto Ricans (**Figure E4**). Moreover, a meta-analysis of both ethnic subgroups did not yield any additional admixture mapping peak at genome-wide significant level.

Fine-mapping of common variants

We next performed fine-mapping of common variants via allelic association testing (**Table 3, Table E2**). We found two regions where genetic variation was significantly associated with severe asthma exacerbations in Mexican Americans after accounting for the number of variants tested within each peak. Within an admixture mapping peak in 5q32, one independent SNP nearby *STK32A* was associated with exacerbations in Mexican Americans (rs28836559, OR for A allele: 1.57, 95% CI: 1.22–2.01, p=3.71x10^{-4}) (**Figure 2A**). Within another 5q32 peak, four independent variants were also associated with severe exacerbations (**Figure 2B**). The minor alleles of SNPs rs1144986 (*C5orf46*) and rs35439318 (*SCGB3A2/CTC-327F10.1*) were associated with lower odds of having exacerbations: OR for G allele: 0.43, 95% CI: 0.28–0.66, p=9.45x10^{-5} and OR for C allele: 0.51, 95% CI: 0.34–0.74, p=5.26x10^{-4}, respectively. Moreover, the minor alleles of rs7704889 (*C5orf46/EEF1GP2*) and rs10035432 (*SCGB3A2/CTC-327F10.1*) were associated with higher risk of exacerbations: OR for A allele: 1.58, 95% CI: 1.23–2.11, p=4.00x10^{-4} and OR for A allele: 1.61, 95% CI: 1.23–2.11, p=4.82x10^{-4}, respectively. No significant associations were detected in any of the regions in Puerto Ricans or at 9q22.2-q22.31 in Mexican Americans (**Table E2)**

The step-wise addition of these SNPs to the regression model that tested the association of each lead ancestry SNP with severe asthma exacerbations revealed that the independent SNPs accounted for the two 5q32 peaks (**Table E3**). Moreover, the association of the SNPs revealed by fine-mapping and severe exacerbations remained significant after additional adjustment by local ancestry and no significant interaction with local ancestry was found (**Table E4**). To assess the robustness of the associations to socio-economic and clinical factors, we performed sensitivity analyses with adjustment by second-hand smoking exposure, insurance status, income quartile,

maternal education, and obesity, confirming that the effects were not explained by these factors (**Table E5**).

We next assessed if these two independent variants replicated among Latino subgroups. In Puerto Ricans, the SNPs rs28836559 (OR for A allele: 1.36, 95% CI: 1.06–1.74, p =1.44x10^{-2}) and rs1144986 (OR for G allele: 0.79, 95% CI: 0.62–1.00, p=4.94x10^{-2}) had consistent effects compared to Mexican Americans. In the meta-analysis of Latinos, the A allele of rs28836559 was associated with higher of having exacerbations (OR: 1.46, 95% CI: 1.23–1.74, p=2.19x10^{-5}, Cochran's Q p=4.22x10^{-1}) and the G allele of rs1144986 showed a protective effect over exacerbations (OR: 0.60, 95% CI: 0.33–1.08, p=1.61x10^{-4}, Cochran's Q p=1.40x10^{-2}). Moreover, an analysis in 1,462 Mexican Americans and 2,346 Puerto Ricans revealed that these SNPs were not associated with the underlying asthma susceptibility at p<0.05 (**Table E6**).

The two SNPs that showed significant and consistent effects in Latino subgroups were assessed for validation in non-Latino populations. However, no association was found with severe exacerbations (p>0.05) in Singaporean Chinese, African Americans, or Spanish individuals, or with less severe exacerbations, including school absences in 421 Swedish children or wheezing in the last year in 100 Spanish individuals (**Table E7**).

Fine-mapping of rare variants

For those regions where no common variant accounted for the admixture mapping peak, we tested the association of rare variants with severe asthma exacerbations. In Mexican Americans, we found 3 significant windows at peak-wise q<0.05 in the 5:148492902-148738466 peak: 5:148503902-148504902 (n_{SNPs}=13, p=2.08x10^{-4}, q=4.02x10^{-2}); 5:148504402-148505402 (n_{SNPs}=20, p=5.30x10^{-4}, q=4.61x10^{-2}); 5:148497902-148498902 (n_{SNPs}=14, p=7.18x10^{-4}, q=4.61x10^{-2}). These variants, located in an intronic region of *HTR4*, were identified to have a "modifier" impact based on their genomic locations and coding effects. A drop-one-out analysis revealed a similar contribution of each individual variant to the association of the regions with severe asthma exacerbations (**Table E8**). The association remained robust to additional adjustment by the covariates described above (**Table E9**). However, these three regions (±25kb) did not replicate in Puerto Ricans or African Americans at q<0.05. Moreover, no significant associations were found within the admixture mapping peaks identified in Puerto Ricans (q>0.05).

Assessment of functional effects

We next conducted QTL mapping for two independent SNPs that replicated Hispanic/Latino subgroups. The effect of rs1144986 and rs28836559 on DNA methylation in whole blood was evaluated for 196 and 468 CpG sites located within ±1 Megabase (Mb) of those SNPs, respectively. The SNPs rs1144986 and rs28836559 were identified as meQTL at q<0.05 (**Table 4**). While the SNP rs28836559 modified methylation levels at the probe cg17608492, located at 10.3 kb of *DPYSL3* (p=1.22x10^{-5}, q=4.02x10^{-3}), the SNP rs1144986 was significantly associated with

three probes annotated to *DPYSL3,* being cg04833034, at 57.4 kb of *DPYSL3,* the most significant CpG (p=5.17x10^{-6}, q=3.41x10^{-3}). Notably, the pair rs1144986-cg10930901 was replicated in the GoDMC results[25] (p= 7.10x10^{-29}). *DPYSL3* expression in bronchial epithelial cells was found to be increased in severe asthma in publicly available gene expression datasets (**Figure E5**). Additionally, the meQTL rs1144986 showed evidence of chromatin interaction with *C5orf46* in lymphoblastoid cells (CHiCAGO score=11.7) and association with H3K4me1 histone marks in blood (**Table E4**). Moreover, rs28836559 was found to significantly interact with *STK32A* in CD34 cells (CHiCAGO score=10.8) and lymphoblastoid cells (CHiCAGO score=10.9). Notably, none of the significant CpGs regulated the expression of nearby genes in whole blood in a subset of 126 Puerto Ricans and 40 Mexicans with DNA methylation and gene expression data available at a false discovery rate-adjusted p<0.05 (**Table E10**).

The effect of the two variants over the expression of genes with transcription start site within ±1Mb was assessed. From the 10 SNP-gene pairs, the SNP rs1144986 showed significant association with gene expression of *SCGB3A2* at q<0.05 in whole-blood (p=6.62x10^{-22}; q=6.62x10^{-20}) and the association was also validated in blood according to the GTEx (p=5.1x10^{-10})[23] and Phenoscanner (minimum p=4.3x10^{-11}) databases.[24]

DISCUSSION

To the best of our knowledge, this study is the first admixture mapping analysis of severe asthma exacerbations in Hispanic/Latino children with asthma. Briefly, we identified several genomic regions in which local Indigenous American or European ancestry was associated with severe asthma exacerbations in Puerto Ricans and Mexican Americans. Although fine-mapping in Puerto Ricans revealed no significant variants, we found five independent SNPs explaining the admixture mapping peak for Indigenous American ancestry at chromosome 5q32 in Mexican Americans, including the SNPs rs28836559 (*STK32A*) and rs1144986 (*C5orf46*), which showed significant consistent effects among Hispanic/Latino subgroups. Interestingly, none of the five loci were associated with asthma exacerbations in non-Latino populations, suggesting ethnic-specific effects shared among Hispanic/Latino subgroups. This could be explained, at least partially, due to differences in allele frequency and ethnic background. For example, the effect allele (A) of rs28836559 is more common in admixed populations (MAF=0.41, 0.25, 0.30 in Mexican Americans, Puerto Ricans, and African Americans, respectively) than in other ethnic groups (**Tables E4, E6**). The fact that Indigenous American ancestry is around 2% in African Americans[28] could contribute to the lack of validation of these genetic variants despite the higher MAF observed.

An analysis of functional effects revealed that the two variants replicated in Puerto Ricans altered DNA methylation nearby *DPYSL3*. The protein DPYSL3 is involved in cytoskeleton remodeling by

participating in the signaling of class 3 semaphorins. While the role of *DPYSL3* in asthma is unclear, it is co-expressed with genes involved in airway type 2 inflammation[29]. However, *DPYSL3* expression in whole blood data from Hispanics/Latinos could not be evaluated due to low/lack expression, consistently with the GTEx data.[23] Additionally, the protective allele of rs1144986 increased *SCGB3A2* expression in blood (**Figure 3**), although *SCGB3A2* is predominantly expressed in the lung.[23] *SCGB3A2* is involved in lung development[30] and has shown an anti-inflammatory role in mice.[31–34] Lung expression levels of pro-inflammatory cytokines, IL-4, IL-5, and IL-13, are lower in ovalbumin-induced mice pre-treated with SCGB3A2 compared to those not treated with it.[33,34] Plasma *SCGB3A2* levels are decreased in severe asthma,[35] and *SCGB3A2* polymorphisms have been associated with asthma susceptibility.[36] In human bronchial epithelial cells, SCGB3A2 decreased airway inflammation inhibiting ERK and JNK activation.[37] In lung cancer, SCGB3A2 induced pyroptotic cell death. Interestingly, patients with higher *SCGB3A2* gene expression manifested higher survival rates.[38] Moreover, *SCGB3A2* shows anti-fibrotic activity through increased expression of STAT1 phosphorylation and *SMAD7* expression,[39,40] hence, decreasing SMAD2/3 phosphorylation, which attenuates the TGFβ signaling,[40] a key pathway implicated in airway remodeling,[41] allergic airway inflammation,[42] and drug response in asthma.[43] Taken together, this evidence supports that these genes may be involved in susceptibility to asthma exacerbations and merit further exploration to understand their specific role.

Moreover, we also found three regions in an intronic region of *HTR4* where rare variation was associated with severe asthma exacerbations in Mexican Americans, but these were not significant in the other populations. Genetic variation of *HTR4* has been associated with lung function and airway obstruction, which could be explained due to the involvement of *HTR4* in lung development.[44] However, the association *HTR4* with severe exacerbations was not confounded by lung function, and these regions were not associated with forced expiratory volume in one second, forced vital capacity, or their ratio in Mexican Americans (**Table E9**).

Our study has several strengths and limitations. We focused on minority populations at high risk of asthma that have undergone distinct historical processes and show differential asthma exacerbations rates. Although we identified several regions in which local ancestry was associated with severe exacerbations, analyzing two different Hispanic/Latino subgroups may have hindered our power to detect associations. Specifically, this could have affected the identification of SNPs in the regions where Indigenous American ancestry was associated with exacerbations in Puerto Ricans, as this ancestral population has a smaller contribution to their genetic background compared to Mexican Americans. Moreover, we performed fine mapping evaluating both common and rare genetic variants. Although genetic signals identified by fine-mapping exceeded the respective peak-wise threshold of significance (**Table 2, Table E2**), none would have exceeded a stringent correction for the total number of variants tested in the discovery stage (p=0.05/13,905 SNPs=3,59x10^{-6}). Furthermore, we also sought to determine if these genetic variants may exert

the same effects in non-Hispanic/Latino populations. The fact that none of the variants was validated in non-Hispanics/Latinos suggests that they could exhibit ethnic-specific effects. We attempted to overcome these limitations by performing functional analyses to reveal the effects of the variants identified on gene expression and DNA methylation.

In summary, we leveraged local ancestry to identify genomic regions that contribute to severe asthma exacerbations in Latinos. Indigenous American ancestry was associated with asthma exacerbation risk at 5q32-q33.1 and novel associations of genetic variants with severe asthma exacerbations with a potential population-specific effect were uncovered. Moreover, these variants had functional effects on *SCGB3A2* gene expression and *DPYSL3* DNA methylation, two genes that are plausibly implicated in severe exacerbations.

ACKNOWLEDGEMENTS

We acknowledge the patients, families, recruiters, health care providers and community clinics for their participation and the contribution of the high-performance compute cluster Wynton HPC underlying UCSF's Research Computing Capability to the results of this research. We also thank Sandra Salazar for her support as GALA II and SAGE study coordinator and the New York Genome Center investigators and teams for whole-genome sequencing sample preparation, quality control, data generation, data processing, and initial joint genotyping. We gratefully acknowledge the studies and participants who provided biological samples and data for TOPMed. We also gratefully acknowledge the contributions of the investigators of the NHLBI TOPMed Consortium (**https://www.nhlbiwgs.org/topmed-banner-authorship**). The genotyping of the GEMAS and MEGA studies was performed at the Spanish National Cancer Research Centre, in the Human Genotyping lab, a member of CeGen, PRB3, and is supported by grant PT17/0019, of the PE I+D+i 2013-2016, funded by ISCIII and ERDF. ISGlobal acknowledges support from the Spanish Ministry of Science and Innovation through the "Centro de Excelencia Severo Ochoa 2019-2023" Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program.

AVAILABILITY OF DATA AND MATERIALS

TOPMed WGS and RNA-seq data from GALA II are available on dbGaP under accession number phs000920.v4.p2. TOPMed WGS data from SAGE is available on dbGaP under accession number phs000921.v4.p1.

ROLE OF THE FUNDING SOURCE

The funders of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report, or decision to submit the report for publication.

REFERENCES

1 Dharmage SC, Perret JL, Custovic A. Epidemiology of Asthma in Children and Adults. *Front Pediatr* 2019;**7**:246. doi:10.3389/fped.2019.00246

2 Hernandez-Pacheco N, Flores C, Oh SS, *et al.* What Ancestry Can Tell Us About the Genetic Origins of Inter-Ethnic Differences in Asthma Expression. *Curr Allergy Asthma Rep* 2016;**16**:53. doi:10.1007/s11882-016-0635-4

3 National Health Interview Survey: Most Recent Asthma Data. Centers for Disease Control and Prevention. 2019. Available at https://www.cdc.gov/asthma/most_recent_data.html (Access date: June 2021).

4 Oraka E, Iqbal S, Flanders WD, *et al.* Racial and ethnic disparities in current asthma and emergency department visits: findings from the National Health Interview Survey, 2001-2010. *J Asthma* 2013;**50**:488–96. doi:10.3109/02770903.2013.790417

5 Nurmagambetov T, Kuwahara R, Garbe P. The Economic Burden of Asthma in the United States, 2008-2013. *Ann Am Thorac Soc* 2018;**15**:348–56. doi:10.1513/AnnalsATS.201703-259OC

6 Krishnan V, Diette GB, Rand CS, *et al.* Mortality in patients hospitalized for asthma exacerbations in the United States. *Am J Respir Crit Care Med* 2006;**174**:633–8. doi:10.1164/rccm.200601-007OC

7 Briggs A, Nasser S, Hammerby E, *et al.* The impact of moderate and severe asthma exacerbations on quality of life: a post hoc analysis of randomised controlled trial data. *J patient-reported outcomes* 2021;**5**:6. doi:10.1186/s41687-020-00274-x

8 Calhoun WJ, Haselkorn T, Miller DP, *et al.* Asthma exacerbations and lung function in patients with severe or difficult-to-treat asthma. *J Allergy Clin Immunol* 2015;**136**:1125-7.e4. doi:10.1016/j.jaci.2015.05.014

9 Yan Q, Forno E, Herrera-Luis E, *et al.* A genome-wide association study of severe asthma exacerbations in Latino children and adolescents. *Eur Respir J* 2021;**57**:2002693. doi:10.1183/13993003.02693-2020

10 Hernandez-Pacheco N, Farzan N, Francis B, *et al.* Genome-wide association study of inhaled corticosteroid response in admixed children with asthma. *Clin Exp Allergy* 2019;**49**:789–98. doi:10.1111/cea.13354

11 Herrera-Luis E, Espuela-Ortiz A, Lorenzo-Diaz F, *et al.* Genome-wide association study reveals a novel locus for asthma with severe exacerbations in diverse populations. *Pediatr Allergy Immunol* 2021;**32**:106–15. doi:10.1111/pai.13337

12 Galanter JM, Gignoux CR, Torgerson DG, *et al.* Genome-wide association study and admixture mapping identify different asthma-associated loci in Latinos: the Genes-environments & Admixture in Latino Americans study. *J Allergy Clin Immunol* 2014;**134**:295–305. doi:10.1016/j.jaci.2013.08.055

13 Gignoux CR, Torgerson DG, Pino-Yanes M, *et al.* An admixture mapping meta-analysis implicates genetic variation at 18q21 with asthma susceptibility in Latinos. *J Allergy Clin Immunol* 2019;**143**:957–69. doi:10.1016/j.jaci.2016.08.057

14 Pino-Yanes M, Gignoux CR, Galanter JM, *et al.* Genome-wide association study and admixture mapping reveal new loci associated with total IgE levels in Latinos. *J Allergy Clin Immunol* 2015;**135**:1502–10. doi:10.1016/j.jaci.2014.10.033

15 Goddard PC, Keys KL, Mak ACY, *et al.* Integrative genomic analysis in African American children with asthma finds three novel loci associated with lung function. *Genet Epidemiol* 2021;**45**:190–208. doi:10.1002/gepi.22365

16 Nishimura KK, Galanter JM, Roth LA, *et al.* Early-life air pollution and asthma risk in minority children. The GALA II and SAGE II studies. *Am J Respir Crit Care Med* 2013;**188**:309–18.

doi:10.1164/rccm.201302-0264OC

17 Reddel HK, Taylor DR, Bateman ED, *et al.* Asthma control and exacerbations - Standardizing endpoints for clinical asthma trials and clinical practice. *Am J Respir Crit Care Med* 2009;**180**:59–99. doi:10.1164/rccm.200801-060ST

18 U.S. Department of Health and Human Services NI of H. *Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma*. National Institute of Health's National Heart Lung and Blood Institute 2007.

19 Pino-Yanes M, Thakur N, Gignoux CR, *et al.* Genetic ancestry influences asthma susceptibility and lung function among Latinos. *J Allergy Clin Immunol* 2015;**135**:228–35. doi:10.1016/j.jaci.2014.07.053

20 Maples BK, Gravel S, Kenny EE, *et al.* RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. *Am J Hum Genet* 2013;**93**:278–88. doi:10.1016/j.ajhg.2013.06.020

21 Lee S, Wu MC, Lin X. Optimal tests for rare variant effects in sequencing association studies. *Biostatistics* 2012;**13**:762–75. doi:10.1093/biostatistics/kxs014

22 Stegle O, Parts L, Durbin R, *et al.* A Bayesian framework to account for complex non-genetic factors in gene expression levels greatly increases power in eQTL studies. *PLoS Comput Biol* 2010;**6**:e1000770. doi:10.1371/journal.pcbi.1000770

23 GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* 2020;**369**:1318–30. doi:10.1126/science.aaz1776

24 Kamat MA, Blackshaw JA, Young R, *et al.* PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* 2019;**35**:4851–3. doi:10.1093/bioinformatics/btz469

25 Min JL, Hemani G, Hannon E, *et al.* Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat Genet* 2021;**53**:1311–21. doi:10.1038/s41588-021-00923-x

26 Schofield EC, Carver T, Achuthan P, *et al.* CHiCP: a web-based tool for the integrative and interactive visualization of promoter capture Hi-C datasets. *Bioinformatics* 2016;**32**:2511–3. doi:10.1093/bioinformatics/btw173

27 Shumyatcher M, Hong R, Levin J, *et al.* Disease-Specific Integration of Omics Data to Guide Functional Validation of Genetic Associations. *AMIA . Annu Symp proceedings AMIA Symp* 2017;**2017**:1589–96.http://www.ncbi.nlm.nih.gov/pubmed/29854229

28 Bryc K, Durand EY, Macpherson JM, *et al.* The genetic ancestry of African Americans, Latinos, and European Americans across the United States. *Am J Hum Genet* 2015;**96**:37–53. doi:10.1016/j.ajhg.2014.11.010

29 Modena BD, Bleecker ER, Busse WW, *et al.* Gene Expression Correlated with Severe Asthma Characteristics Reveals Heterogeneous Mechanisms of Severe Disease. *Am J Respir Crit Care Med* 2017;**195**:1449–63. doi:10.1164/rccm.201607-1407OC

30 Kurotani R, Tomita T, Yang Q, *et al.* Role of secretoglobin 3A2 in lung development. *Am J Respir Crit Care Med* 2008;**178**:389–98. doi:10.1164/rccm.200707-1104OC

31 Kido T, Yoneda M, Cai Y, *et al.* Secretoglobin superfamily protein SCGB3A2 deficiency potentiates ovalbumin-induced allergic pulmonary inflammation. *Mediators Inflamm* 2014;**2014**. doi:10.1155/2014/216465

32 Yoneda M, Xu L, Kajiyama H, *et al.* Secretoglobin Superfamily Protein SCGB3A2 Alleviates House Dust Mite-Induced Allergic Airway Inflammation in Mice. *Int Arch Allergy Immunol* 2016;**171**:36–44. doi:10.1159/000450788

33 Chiba Y, Kurotani R, Kusakabe T, *et al.* Uteroglobin-related protein 1 expression suppresses allergic airway inflammation in mice. *Am J Respir Crit Care Med* 2006;**173**:958–64.

doi:10.1164/rccm.200503-456OC

34 Mandal AK, Zhang Z, Ray R, *et al.* Uteroglobin represses allergen-induced inflammatory response by blocking PGD2 receptor-mediated functions. *J Exp Med* 2004;**199**:1317–30. doi:10.1084/jem.20031666

35 Inoue K, Wang X, Saito J, *et al.* Plasma UGRP1 levels associate with promoter G-112A polymorphism and the severity of asthma. *Allergol Int* 2008;**57**:57–64. doi:10.2332/allergolint.O-07-493

36 Zhao G, Lin X, Zhou M, *et al.* Association between CC10 +38A/G polymorphism and asthma risk: A meta-analysis. *Pakistan J Med Sci* 2013;**29**:1439–43. doi:10.12669/pjms.296.3724

37 Wang X, Tanino Y, Sato S, *et al.* Secretoglobin 3A2 attenuates lipopolysaccharide-induced inflammation through inhibition of ERK and JNK pathways in bronchial epithelial cells. *Inflammation* 2015;**38**:828–34. doi:10.1007/s10753-014-9992-0

38 Yokoyama S, Nakayama S, Xu L, *et al.* Secretoglobin 3A2 eliminates human cancer cells through pyroptosis. *Cell death Discov* 2021;**7**:12. doi:10.1038/s41420-020-00385-w

39 Cai Y, Kimura S. Secretoglobin 3A2 Exhibits Anti-Fibrotic Activity in Bleomycin-Induced Pulmonary Fibrosis Model Mice. *PLoS One* 2015;**10**:e0142497. doi:10.1371/journal.pone.0142497

40 Kurotani R, Okumura S, Matsubara T, *et al.* Secretoglobin 3A2 suppresses bleomycin-induced pulmonary fibrosis by transforming growth factor beta signaling down-regulation. *J Biol Chem* 2011;**286**:19682–92. doi:10.1074/jbc.M111.239046

41 Sagara H, Okada T, Okumura K, *et al.* Activation of TGF-beta/Smad2 signaling is associated with airway remodeling in asthma. *J Allergy Clin Immunol* 2002;**110**:249–54. doi:10.1067/mai.2002.126078

42 Hansen G, McIntire JJ, Yeung VP, *et al.* CD4(+) T helper cells engineered to produce latent TGF-beta1 reverse allergen-induced airway hyperreactivity and inflammation. *J Clin Invest* 2000;**105**:61–70. doi:10.1172/JCI7589

43 Burchard EG, Avila PC, Nazario S, *et al.* Lower bronchodilator responsiveness in Puerto Rican than in Mexican subjects with asthma. *Am J Respir Crit Care Med* 2004;**169**:386–92. doi:10.1164/rccm.200309-1293OC

44 Hodge E, Nelson CP, Miller S, *et al.* HTR4 gene structure and altered expression in the developing lung. *Respir Res* 2013;**14**:1–13. doi:10.1186/1465-9921-14-77

Table 1. Characteristics of the asthma patients recruited between 2006 and 2014 for the Genes-Environments and Admixture in Latino Asthmatics (GALA II) study.

	Puerto Ricans (N=1,124)			Mexican Americans (N=625)		
	Exacerbators	Non-exacerbators	*p*-value	Exacerbators	Non-exacerbators	*p*-value
Number of individuals	820	304		223	402	
Sex (% male)	443 (54.0)	160 (52.6)	7.27×10^{-1}	129 (57.8)	226 (56.2)	7.57×10^{-1}
Age, mean+SD (years)	12.2±3.2	13.8±3.6	1.83×10^{-11}	12.2±3.2	13.2±3.4	6.16×10^{-4}
European ancestry, mean±SD (%)	62.6±10.7	63.2±10.5	5.45×10^{-1}	38.6±14.7	40.0±13.5	3.03×10^{-1}
African ancestry, mean±SD (%)	23.5±11.7	26.7±11.4	1.57×10^{-1}	4.6±2.2	5.0±3.3	3.03×10^{-1}
Indigenous American ancestry, mean±SD (%)	13.8±3.93	14.3±4.8	1.16×10^{-1}	56.7±15.2	55.0±14.2	2.00×10^{-1}
Treatment category, n (%)						
Step 1	353 (43.0)	227 (74.7)	8.27×10^{-21}	53 (23.7)	185 (46.0)	6.57×10^{-8}
Step 2	239 (29.1)	41 (13.5)	1.07×10^{-7}	86 (38.6)	131 (32.6)	1.57×10^{-1}
Step 3	228 (27.9)	36 (11.8)	3.23×10^{-8}	84 (37.7)	86 (21.4)	1.82×10^{-5}

Abbreviation: SD: Standard deviation. For continuous variables, the mean and standard deviation are displayed and Mann-Whitney-Wilcoxon test was applied for the comparison of exacerbators versus non-exacerbators. For categorical variables, the number and proportion of subjects in each category are shown and a χ2 test was applied for the comparison of exacerbators versus non-exacerbators.

Table 2. Significant admixture mapping peaks identified in Mexican Americans and Puerto Ricans.

Ethnicity/Ancestry	Chr. band	Start (Bp)[a]	End (Bp)[a]	Peak size (Kb)	rsID[b]	OR (95%CI)	*p*-value
Puerto Rican/ Indigenous American	3p13	72283579	72585953	302	rs4677148	0.57 (0.44-0.74)	2.55×10^{-5}
	3p13	74273323	74349218	76	rs4677382	0.57 (0.43-0.75)	8.57×10^{-5}
	13q13.2	34112347	36450444	2338	rs10514839	2.18 (1.55-3.06)	6.56×10^{-6}
Mexican American/ Indigenous American	5q32	147235698	147305729	70	rs72822024	1.66 (1.28-2.14)	1.01×10^{-4}
	5q32	147809331	147911637	102	rs10477350	1.73 (1.34-2.24)	3.23×10^{-5}
	5q32	148492902	148738466	246	rs34729481	1.68 (1.3-2.16)	7.60×10^{-5}
Mexican American/ European	5q33.1	151144426	151181737	37	rs884083	0.59 (0.45-0.77)	1.25×10^{-4}
	9q22.2-q22.31	91139503	91737237	598	rs7034207	0.58 (0.44-0.75)	5.65×10^{-5}
	9q22.31	92163398	93013279	850	rs11794664	0.58 (0.45-0.76)	7.41×10^{-5}

[a]Coordinates are referred to the human reference genome assembly GRCh38. [b]Genetic variant with the minimum association p-value within the peak.

Abbreviations: Bp: Base pairs; Chr: Chromosome; CI: confidence interval; Kb: Kilobases; OR: Odds ratio; rsID: reference single nucleotide polymorphism identifier; iSNPs: independent single nucleotide polymorphisms; U95CI: Upper limit of the 95% confidence interval.

Table 3. Association results of independent genetic variants at chromosome 5 that were peak-wise significantly associated with severe asthma exacerbations.

				Mexican Americans			Puerto Ricans			Meta-analysis	
rsID	Position (Bp)[a]	A1/A2	Closest gene	EAF	OR (95% CI)	*p*-value	EAF	OR (95% CI)	*p*-value	OR (95%CI)	*p*-value
rs28836559	147294667	A/G	*STK32A*	0.41	1.57 (1.22–2.01)	**3.71×10^{-4}**	0.25	1.36 (1.06–1.74)	**1.44×10^{-2}**	1.46 (1.23–1.74)	**2.20×10^{-5}**
rs1144986	147883415	G/A	*C5orf46*	0.13	0.43 (0.28–0.66)	**9.45×10^{-5}**	0.22	0.79 (0.62–1.00)	**4.94×10^{-2}**	0.60 (0.33–1.08)	**1.61×10^{-4}** [b]
rs7704889	147911637	A/G	*C5orf46/EEF1GP2*	0.54	1.58 (1.23–2.03)	**4.00×10^{-4}**	0.42	0.97 (0.79–1.20)	7.94×10^{-1}	1.23 (0.77–1.98)	**6.57×10^{-3}** [b]
rs10035432	147855193	A/G	*SCGB3A2/CTC-327F10.1*	0.26	1.61 (1.23–2.11)	**4.82×10^{-4}**	0.21	1.02 (0.81–1.29)	8.72×10^{-1}	1.28 (0.81–2.00)	**5.93×10^{-3}** [b]
rs35439318	147862111	C/T	*SCGB3A2/CTC-327F10.1*	0.15	0.51 (0.34–0.74)	**5.26×10^{-4}**	0.22	0.83 (0.65–1.05)	1.27×10^{-1}	0.66 (0.41–1.08)	**1.45×10^{-3}** [b]

[a]Coordinates are referred to the human reference genome assembly GRCh38; [b]A random effect model was applied since heterogeneity was found between populations (Cochran's Q p-value<0.05).

Association p-values<0.05 are shown in bold. Abbreviations: A1: Effect allele; A2: Non-effect allele; Bp: Base pairs; CI: Confidence interval; EAF: Effect allele frequency; OR: Odds ratio; rsID: reference single nucleotide polymorphism identifier.

Table 4. Significant results from the QTL mapping in Latinos.

				Puerto Ricans			**Mexican Americans**			**Meta-analysis**[a]		
SNP - analysis	Target	Position[b]	Gene	Coef (SE)	*p*-value	PQ	Coef (SE)	*p*-value	PQ	Coef (SE)	*p*-value	Q
	cg04833034	147452619	*DPYSL3*	0.18 (0.04)	**$1.58x10^{-5}$**	0.66	0.11 (0.07)	$1.28x10^{-1}$	NA[c]	0.16 (0.04)	**$6.36x10^{-6}$**	**$3.55x10^{-3}$**
rs1144986 -meQTL	cg09639133	147451646	*DPYSL3*	0.22 (0.05)	**$5.10x10^{-5}$**	0.53	0.14 (0.07)	$5.25x10^{-2}$	NA[c]	0.19 (0.04)	**$1.13x10^{-5}$**	**$3.55x10^{-3}$**
	cg10930901	147505871	*DPYSL3*	-0.18 (0.05)	**$7.96x10^{-5}$**	0.26	-0.05 (0.07)	$4.56x10^{-1}$	NA[c]	-0.14 (0.04)	**$1.84x10^{-4}$**	**$2.88x10^{-2}$**
rs1144986 - eQTL	ENSG00000164265	147870682	*SCGB3A2*	0.68 (0.08)	**$1.01x10^{-15}$**	NA	0.47 (0.10)	**$1.85x10^{-6}$**	NA	0.60 (0.06)	**$6.12x10^{-21}$**	**$8.57x10^{-20}$**
rs28836559 - meQTL	cg17608492	147520349	*DPYSL3*	-0.14 (0.04)	**$3.77x10^{-4}$**	NA[d]	-0.22 (0.09)	**$2.23x10^{-2}$**	NA[c]	-0.15 (0.03)	**$1.96x10^{-5}$**	**$4.08x10^{-3}$**

[a]A fixed effect model was applied since no heterogeneity was found between populations (Cochran's Q p-value>0.05). [b]Chromosomal positions referred to the genome assembly GRCh38. For CpG sites, the location of the CpG is stated. For genes, the location of the transcription start site is shown. [c]The SNP did not meet the genotype inclusion criteria for the individuals profiled with the EPIC array. [d]The CpG was not available in the 450K array.
Association p-values<0.05 are shown in bold
Abbreviations PQ: Cochran's Q-test p-value; Coef: Regression coefficient estimate; NA: Not available/applicable; Q: Q-value; SE: Standard error of the regression coefficient estimate.

FIGURES

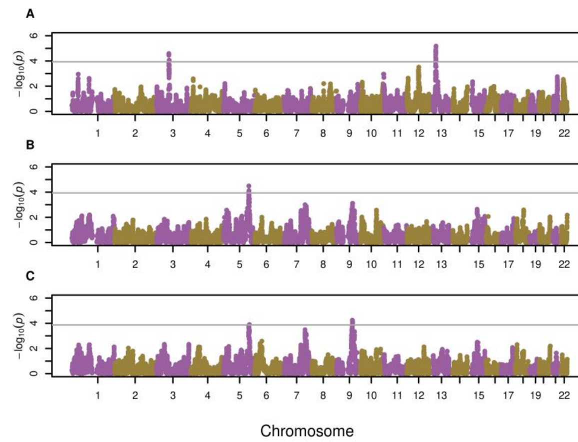


Figure 1. Manhattan plot for the admixture mapping results of severe exacerbations. A) Association results for Indigenous American ancestry in Puerto Ricans B) Association results for Indigenous American ancestry in Mexican Americans C) Association results for European ancestry in Mexican Americans. The association with local ancestry is represented as -$\log_{10}$ p-value on the *y*-axis along the chromosomes (*x*-axis). The suggestive significance threshold for replication is indicated by the grey line.

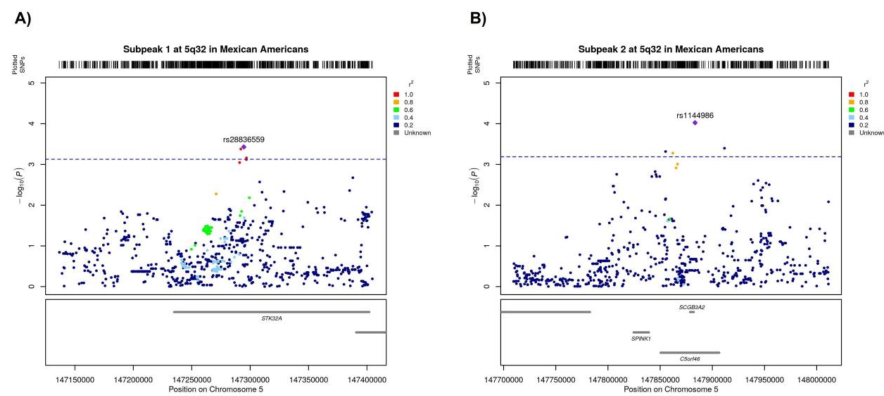


Figure 2. Locus zoom for the fine-mapping results displaying the region around the most significant signals at 5q32. The statistical significance of association results (-$\log_{10}$ p-value) is represented for each SNP as a dot (left *y*-axis) by chromosome position (*x*-axis). SNPs are color-coded to show their linkage disequilibrium (LD) with the most significant SNP based on the pairwise r^2 values from Mexican Americans. The peak-wise significance threshold is represented as a dashed blue line.

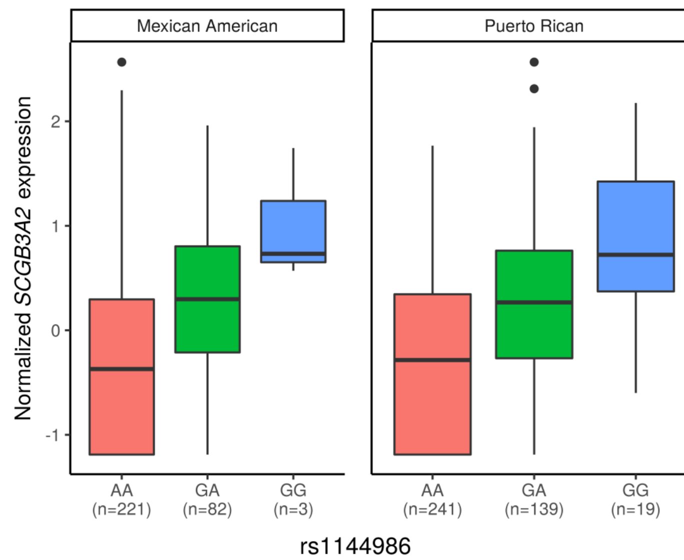


Figure 3. Gene expression levels of *SCGB3A2* in whole blood by genotype at rs1144986 in Mexican American and Puerto Rican asthma patients.

Link to Supporting Information:

https://drive.google.com/drive/folders/1IZTRl7NtQiWPAik-0ygUMGUHxBQFBwjA?usp=sharing

4.5 Multi-ancestry genome-wide association study of asthma exacerbations

This *Chapter* reports the results of the first multi-ethnic GWAS of AEs in patients with asthma, including European descents, Singaporean Chinese, Latinos and African Americans. Given that the genetic variation of AEs has been explored separately in populations of different ethnic background, the aim of this study was to identify novel genetic associations for AEs that are shared among different ancestral backgrounds. The GWAS design consisted in a multi-ethnic discovery stage to identify suggestive signals of association, and a replication stage to validate these findings and a meta-analysis of both stages. Asthma exacerbations were defined based on the report of acute asthma care, oral corticosteroids use, asthma-related hospitalizations, and/or school absences due to asthma. The functional effects of the genetic variants that showed significant and consistent effect across the different stages were assessed via meQTL analysis in blood from Hispanics/Latinos and African Americans and *in silico* analysis.

Two genetic variants showed significant and consistent effect for the association with AEs. These variants were located within the pantothenate kinase 1 (*PANK1*) gene and the intergenic region of the vascular cell adhesion molecule 1 (*VCAM1*) and exostosin-like 2 (*EXTL2*) gene. The meQTL analysis correlated genotype dosage and DNA methylation at several nearby CpG sites in whole-blood from Hispanics/Latinos and African Americans with asthma. Moreover, *in silico* analysis revealed that gene expression of *EXTL2* and *PANK1* was regulated by these loci. In summary, our study revealed two functional genetic variants within genomic regions containing potential targets of interest whose specific role in AEs should be further explored.

This *Chapter* is presented as a manuscript draft entitled 'Multi-ancestry genome-wide association study of asthma exacerbations' that will be submitted to an international scientific journal to be considered for publication.

Multi-ancestry genome-wide association study of asthma exacerbations

Running title: Multi-ancestry GWAS of asthma exacerbations

AUTHORS

Esther Herrera-Luis[1], Victor E. Ortega[2], Elizabeth J. Ampleford[3], Yang Yie Sio[4], Raquel Granell[5], Emmely de Roos[6,7], Natalie Terzikhan[6,7] , Ernesto Elorduy Vergara[8], Natalia Hernandez-Pacheco[9,10], Javier Perez-Garcia[1], Elena Martin-Gonzalez[1], Fabian Lorenzo-Diaz[1,11], Simone Hashimoto[12], Paul Brinkman[12], U-BIOPRED Study Group[#], Andrea L. Jorgensen[13], Qi Yan[14], Erick Forno[15], Susanne J. Vijverberg[16,17,18], Ryan Lethem[5], Antonio Espuela-Ortiz[1], Mario Gorenjak[19], Celeste Eng[20], Ruperto González-Pérez[21,22], José M Hernández-Pérez[23,24], Paloma Poza-Guedes[21,22], Olaia Sardón[25,26], Paula Corcuera[25], Greg A. Hawkins[27], Annalisa Marsico[28], Thomas Bahmer[29], Klaus F. Rabe[29], Gesine Hansen[30], Matthias Volkmar Kopp[31,32,33], Raimon Rios[34], Maria Jesus Cruz[10,35], Francisco-Javier González-Barcala[36], José María Olaguibel[10,37], Vicente Plaza[10,38], Santiago Quirce[10,39], Glorisa Canino[40], Michelle Cloutier[41], Victoria del Pozo[10,42], Jose R Rodriguez-Santana[43], Javier Korta-Murua[26], Jesús Villar[10,44], Uroš Potočnik[45], Camila Figueiredo[46], Michael Kabesch[47], Somnath Mukhopadhyay[48,49], Munir Pirmohamed[50], Daniel B. Hawcutt[51,52,53], Erik Melén[9,54], Colin N Palmer[49], Steve Turner[55], Anke H. Maitland-van der Zee[16,17,18], Erika von Mutius[56,57,58], Juan C. Celedón[15], Guy Brusselle[6,7,59], Fook Tim Chew[4], Eugene Bleecker[60], Deborah Meyers[60], Esteban G Burchard[22,61], Maria Pino-Yanes[1,10,62]

From:

[1]Genomics and Health Group, Department of Biochemistry, Microbiology, Cell Biology and Genetics, Universidad de La Laguna (ULL), San Cristóbal de La Laguna, Tenerife, Spain;
[2]Department of Internal Medicine, Division of Respiratory Medicine, Mayo Clinic, Scottsdale, Arizona, U.S.A;
[3]Department of Internal Medicine, Center for Precision Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina, U.S.A.;
[4]Department of Biological Sciences, National University of Singapore, Singapore;
[5]MRC Integrative Epidemiology Unit (IEU), Population Health Sciences, Bristol Medical School, University of Bristol, Oakfield Grove, Bristol, U.K.;
[6]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands;
[7]Department of Respiratory Medicine, Ghent University Hospital, Ghent, Belgium;
[8]Institute of Computation Biology, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich, Germany;
[9]Department of Clinical Sciences and Education, Södersjukhuset, Karolinska Institutet, Stockholm, Sweden;
[10]CIBER de Enfermedades Respiratorias (CIBERES), Madrid, Spain;
[11]Instituto Universitario de Enfermedades Tropicales y Salud Pública de Canarias (IUETSPC), Universidad de La Laguna (ULL), San Cristóbal de La Laguna, Tenerife, Spain;
[12]Dept. Respiratory Medicine, Amsterdam University Medical Center, University of Amsterdam, Amsterdam, the Netherlands;
[13]14Department of Health Data Science, Institute of Population Health, University of Liverpool, L69 3GL, Liverpool, U.K.;
[14]Department of Obstetrics and Gynecology, Columbia University Irving Medical Center, New York, New York, U.S.A.;
[15]Division of Pediatric Pulmonary Medicine, UPMC Children's Hospital of Pittsburgh, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.;
[16]Dept of Respiratory Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands;
[17]Division of Pharmacoepidemiology and Clinical Pharmacology, Faculty of Science, Utrecht University, Utrecht, The Netherlands;

[18]Dept of Paediatric Respiratory Medicine and Allergy, Emma's Children Hospital, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands;
[10]Center for Human Molecular Genetics and Pharmacogenomics, Faculty of Medicine, University of Maribor, Maribor, Slovenia.
[20]Department of Medicine, University of California San Francisco, San Francisco, California, U.S.A.;
[21]Allergy Department, Hospital Universitario de Canarias, Santa Cruz de Tenerife, Tenerife, Spain;
[22]Severe Asthma Unit, Allergy Department, Hospital Universitario de Canarias, Santa Cruz de Tenerife, Tenerife, Spain
[23]Pulmonary Medicine, Hospital Universitario de N.S de Candelaria, Santa Cruz de Tenerife, Spain;
[24]Pulmonary Medicine, Hospital General de La Palma, La Palma, Santa Cruz de Tenerife, Spain;
[25]Division of Pediatric Respiratory Medicine, Hospital Universitario Donostia, San Sebastián, Spain;
[26]Department of Pediatrics, University of the Basque Country (UPV/EHU), San Sebastián, Spain;
[27]Department of Biochemistry, Wake Forest School of Medicine, Winston-Salem, North Carolina, U.S.A.;
[28]Computational Health Center, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich, Germany;
[29]LungenClinic Grosshansdorf, Grosshansdorf and Dept of Medicine, Christian Albrechts University Kiel, Germany; Members of the German Center for Lung Research (DZL)
[30]Department of Pediatric Pneumology, Allergology and Neonatology, Hannover Medical School, Hannover, Germany; Members of the German Center for Lung Research (DZL);
[31]Division of Pediatric Pneumology & Allergology, University Medical Center Schleswig-Holstein, Lübeck, Germany;
[32]Airway Research Center North (ARCN), Members of the German Center for Lung Research (DZL);
[33]Department of Paediatric Respiratory Medicine, Inselspital, University Children's Hospital of Bern, University of Bern, Bern, Switzerland;
[34]Programa de Pós Graduação em Imunologia (PPGIm), Instituto de Ciências da Saúde, Universidade Federal da Bahia (UFBA), Salvador, Brazil;
[35]Servicio de Neumología, Hospital Vall d'Hebron, Barcelona, Spain;
[36]Servicio de Neumología, Complejo Hospitalario Universitario de Santiago, Santiago de Compostela, Spain;
[37]Servicio de Alergología, Complejo Hospitalario de Navarra, Pamplona, Navarra, Spain;
[38]Departamento de Medicina Respiratoria, Hospital de la Santa Creu i Sant Pau, Instituto de Investigación Biomédica Sant Pau (IIB Sant Pau), Barcelona, Spain;
[39]Department of Allergy, La Paz University Hospital, IdiPAZ, Madrid, Spain;
[40]Behavioral Sciences Research Institute, University of Puerto Rico, San Juan, Puerto Rico;
[41]Department of Pediatrics, University of Connecticut, Farmington, Connecticut, U.S.A;
[42]Immunology Department, Instituto de Investigación Sanitaria Hospital Universitario Fundación Jiménez Díaz, Madrid, Spain;
[43]Centro de Neumología Pediátrica, San Juan, Puerto Rico;
[44]Multidisciplinary Organ Dysfunction Evaluation Research Network, Research Unit, Hospital Universitario Dr. Negrín, Las Palmas de Gran Canaria, Spain
[45]Laboratory for Biochemistry, Molecular Biology and Genomics, Faculty for Chemistry and Chemical Engineering, University of Maribor, Maribor, Slovenia;
[46]Instituto de Ciências da Saúde, Universidade Federal da Bahia, Salvador, Brazil;
[47]Dept of Paediatric Pneumology and Allergy, University Children's Hospital Regensburg (KUNO), Regensburg, Germany;
[48]Academic Dept of Paediatrics, Brighton and Sussex Medical School, Royal Alexandra Children's Hospital, Brighton, UK;
[49]Population Pharmacogenetics Group, Biomedical Research Institute, Ninewells Hospital and Medical School, University of Dundee, Dundee, UK;
[50]Dept of Pharmacology and Therapeutics, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK;
[51]Dept of Women's and Children's Health, University of Liverpool, Liverpool, UK;
[52]Alder Hey Children's Hospital, Liverpool, UK;
[53]NIHR Alder Hey Clinical Research Facility, Alder Hey Children's Hospital, Liverpool, UK;
[54]Sachs' Children's Hospital, South General Hospital, Stockholm, Sweden;
[55]Child Health, University of Aberdeen, Aberdeen, UK;
[56]Institute for Asthma and Allergy Prevention, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich, Germany;
[57]Dr von Hauner Children's Hospital, LMU München, Munich, Germany;
[58]Comprehensive Pneumology Center Munich (CPC-M), Member of the German Center for Lung Research, Munich, Germany;
[59]Department of Respiratory Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands;;

[60]Department of Internal Medicine, Division of Genetics, Genomics, and Precision Medicine, University of Arizona College of Medicine, Tucson, Arizona, U.S.A.;
[61]Department of Bioengineering and Therapeutic Sciences, University of California San Francisco, San Francisco, California, U.S.A.;
[62]Instituto de Tecnologías Biomédicas (ITB), Universidad de La Laguna (ULL), San Cristóbal de La Laguna, Tenerife, Spain.

#The members of the U-BIOPRED Study Group are provided in the online data supplement.

***Corresponding author:**

Dr. Maria Pino-Yanes

Genomics and Health Group, Department of Biochemistry, Microbiology, Cell Biology and Genetics. Universidad de La Laguna (ULL)

Apartado 456, 38200, San Cristóbal de La Laguna, Tenerife, Spain

Phone: (+34) 922 316502 6343

Fax: (+34) 922 318 490

e-mail: mdelpino@ull.edu.es

Word count: Abstract (247), Main text (3,396)

This manuscript is accompanied by 3 tables, 3 Figures, and a Supporting Information File.

ACKNOWLEDGEMENTS

The authors thank the patients, families, recruiters, health care providers, and community clinics for their participation in the studies analyzed in this manuscript. The authors also acknowledge the contribution of the high-performance compute cluster Wynton HPC underlying UCSF's Research Computing Capability to the results of this research and Sandra Salazar for her support as GALA II and SAGE study coordinator. The authors thank the Centro Nacional de Genotipado-Plataforma de Recursos Biomoleculares-Instituto de Salud Carlos III (CeGen-PRB3-ISCIII; www.cegen.org) for providing the genotyping services. The authors also acknowledge all the families who took part in ALSPAC, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses.

This work was funded by the Spanish Ministry of Science and Innovation MCIN/AEI/10.13039/501100011033, and the European Regional Development Fund "ERDF A way of making Europe" by the European Union (SAF2017-83417R), by MCIN/AEI/10.13039/501100011033 (PID2020-116274RB-I00) and by the Allergopharma-EAACI award 2021. This study was also supported by the SysPharmPedia grant from the ERACoSysMed 1st Joint Transnational Call from the European Union under the Horizon 2020.

GALA II and SAGE studies were supported by the Sandler Family Foundation, the American Asthma Foundation, the RWJF Amos Medical Faculty Development Program, Harry Wm. and Diana V. Hind Distinguished Professor in Pharmaceutical Sciences II, the National Heart, Lung, and Blood Institute of the National Institutes of Health (R01HL117004, R01HL128439, R01HL135156, X01HL134589, R01HL141992, and R01HL141845), National Institute of Health and Environmental Health Sciences (R01ES015794 and R21ES24844); the National Institute on Minority Health and Health Disparities (NIMHD) (P60MD006902, R01MD010443, and R56MD013312); the National Institute of General Medical Sciences (NIGMS) (RL5GM118984); the Tobacco-Related Disease Research Program (24RT-0025 and 27IR-0030); and the National Human Genome Research Institute (NHGRI) (U01HG009080) to EGB. The PACMAN study was funded by a strategic alliance between GlaxoSmithKline and Utrecht Institute for Pharmaceutical Sciences. The SLOVENIA study was financially supported by the Slovenian Research Agency (research core funding No. P3-0067) and from SysPharmPediA grant, co-financed by the Ministry of Education, Science and Sport Slovenia (MIZS) (contract number C3330-16-500106). The SHARE Bioresource (GoSHARE) and SHARE have ongoing funding from NHS Research Scotland and were established by funding from The Wellcome Trust Biomedical Resource [Grant No. 099177/Z/12/Z]. Genotyping of samples from BREATHE, PAGES, and GoSHARE was funded by AC15/00015 and conducted at the Genotyping National Centre (CeGEN) CeGen-PRB3-ISCIII; supported by ISCIII and European Regional Development Fund (ERDF) (PT17/0019). ALSPAC

was supported by the UK Medical Research Council and Wellcome (102215/2/13/2) and the University of Bristol. The Swedish Heart-Lung Foundation, the Swedish Research Council, and Region Stockholm (ALF project and database maintenance) funded the BAMSE study. The PASS study was funded by the NHS Chair of Pharmacogenetics via the UK Department of Health. UBIOPRED was funded by the Innovative Medicines Initiative (IMI) Joint Undertaking, under grant agreement no. 115010, resources for which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and kind contributions from companies in the European Federation of Pharmaceutical Industries and Associations (EFPIA). Genotyping of samples from GEMAS and MEGA studies was funded by the Spanish Ministry of Science and Innovation (SAF2017-87417R) at the Spanish National Cancer Research Centre, in the Human Genotyping lab, a member of CeGen, PRB3, and was supported by grant PT17/0019, of the PE I+D+i 2013-2016, funded by ISCIII and ERDF. The genotyping of GEMAS was also partially funded by Fundación Canaria Instituto de Investigación Sanitaria de Canarias (PIFIISC19/17). The Rotterdam Study was funded by Erasmus Medical Center and Erasmus University Rotterdam; Netherlands Organization for the Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly (RIDE); the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. ALLIANCE Cohort was funded by grants from the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF) as part of the German Centre for Lung Research (DZL) funding. The Hartford-Puerto Rico study was funded by the U.S. National Institutes of Health (grant HL07966 to JCC).

MP-Y was funded by the Ramón y Cajal Program (RYC-2015-17205) by MCIN/AEI/10.13039/501100011033 and by the European Social Fund "ESF Investing in your future". MP-Y and JV were supported by CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Spain (CB/06/06/1088). EH-L was supported by a fellowship awarded by MCIN/AEI/10.13039/501100011033 and by "ESF Investing in your future" (PRE2018-083837).

JP-G was supported by a fellowship awarded by Spanish Ministry of Universities (FPU19/02175). AE-O reports funding from the Spanish Ministry of Science, Innovation, and Universities (MICIU) and Universidad de La Laguna (ULL). NH-P was supported by a Medium-Term Research Fellowship by the European Academy of Allergy and Clinical Immunology (EAACI) and a Long-Term Research Fellowship by the European Respiratory Society (ERS) (LTRF202101-00861).

UP and MG were supported by the Ministry of Education, Science and Sport of the Republic of Slovenia, grant PERMEABLE (contract number C3330-19-252012). SCSGES results were contributed by authors FTC and YYS. FTC has received research support from the Singapore Ministry of Education Academic Research Fund, Singapore Immunology Network (SIgN), National Medical Research Council (NMRC) (Singapore), Biomedical Research Council (BMRC) (Singapore), and the Agency for Science Technology and Research (A*STAR) (Singapore); Grant

Numbers: N-154-000-038-001, R-154-000-191-112, R-154-000-404-112, R-154-000-553-112, R-154-000-565-112, R-154-000-630-112, R-154-000-A08-592, R-154-000-A27-597, R-154-000-A91-592, R-154-000-A95-592, R-154-000-B99-114, BMRC/01/1/21/18/077, BMRC/04/1/21/19/315, SIgN-06-006, SIgN-08-020, NMRC/1150/2008, and H17/01/a0/008. F.T.C. has received consulting fees from Sime Darby Technology Centre; First Resources Ltd; Genting Plantation, and Olam International, outside the submitted work. YYS has received research support from the NUS Resilience & Growth Postdoctoral Fellowships with grant number: R-141-000-036-281. QY conducted the analysis from Hartford-Puerto Rico and United Kingdom Biobank studies. QY was funded by the U.S. National Institutes of Health (HL138098).

All funding agencies had no role in the study design, data collection, and analysis, decision to publish, or preparation of the manuscript. The views expressed are those of the authors and not necessarily those of the any funder, National Health Service (NHS), the National Institute for Health Research (NIHR) or the United Kingdom Department of Health.

CONFLICT OF INTEREST

AE-O received grants from the Spanish Ministry of Science, Innovation, and Universities (MICIU) and Universidad de La Laguna (ULL). EH-L, and MP-Y report funding from the Spanish Ministry of Science and Innovation (MCIN/AEI/10.13039/501100011033) and by “ESF Investing in your future” by the European Union. JP-G reports funding from the Spanish Ministry of Universities. MP-Y and FLD report grants from MCIN/AEI/10.13039/501100011033 and the European Regional Development Fund “ERDF A way of making Europe” by the European Union. MP-Y reports grant support from GlaxoSmithKline, Spain paid to Fundación Canaria Instituto de Investigación Sanitaria de Canarias (FIISC) for a project outside the submitted work. MP-Y and JV reports grants from Instituto de Salud Carlos III, Madrid, Spain. JV also reports funding by ISCIII and the European Regional Development Fund “ERDF A way of making Europe”. JMH-P has received fees from CSL Behring, GSK, Astra-Zeneca, laboratorios Menarini, Boehringer Ingelheim, FAES, laboratorios Esteve, Laboratorios Ferrer, Mundipharma, Laboratorios Rovi, Roche, Novartis, GRIFOLS, Pfizer, Acthelion-Jansen, Chiesi y Laboratorios Bial for the realization of courses, talks, consultancies, and other activities related to his professional activity. FTC has received research support from the Singapore Ministry of Education Academic Research Fund, Singapore Immunology Network (SIgN), National Medical Research Council (NMRC) (Singapore), Biomedical Research Council (BMRC) (Singapore), and the Agency for Science Technology and Research (A*STAR) (Singapore). FTC has received consulting fees from Sime Darby Technology Centre; First Resources Ltd; Genting Plantation, and Olam International, outside the submitted work. YYS has received research support from the NUS Resilience & Growth Postdoctoral Fellowships. UP and MG received grants from the Ministry of Education, Science and Sport from Slovenia, the Slovenian Research Agency. M-JC received grants from the Instituto de Salud Carlos III. DH received grant support from by NIHR for work on NIHR Alder Hey Clinical Research

Facility, received payment for medicolegal report writing not related to asthma or pharmacogenomics for UK family court as an expert in pediatric clinical pharmacology. FJ-B received fees from ALK, Astra-Zeneca (AZ), Bial, Chiesi, Gebro Pharma, GlaxoSmithKline (GSK), Menarini, Rovi, Roxall, Sanofi, Stallergenes-Greer and Teva. G-B received fees from AZ, GSK, Boehringer-Ingelheim, Novartis, Chiesi and Sanofi. JC received research materials from Pharmavite and GSK and Merck in order to provide medications free of cost to participants in NIH-funded studies, unrelated to the current work. VO received grants from the National Heart, Lung, and Blood Institute, has participated in Data Safety Monitoring Boards for Regeneron and Sanofi, and participated as a Chair of the section on Genetics and Genomics of the American Thoracic Society. MVK has received grants from the German Federal Ministry of Education and Research, fees from Allergopharma GmbH, Sanofi Aventis GmbH, Infectopharm GmbH, Vertex GmbH and Leti GmbH, has participated in Data Safety Monitoring Boards for Sanofi Aventis GmbH and is the president of the German-Swiss-Austrian Society of Pediatric Pulmonology (GPP). NHP received support from the Instituto de Salud Carlos III, the European Social Funds from the European Union "ESF invests in your future", the European Academy of Allergy and Clinical Immunology and the European Respiratory Society. MP has received grants from NHS Chair of Pharmacogenetic grant from UK Department of Health, has received partnership funding for the following: MRC Clinical Pharmacology Training Scheme (co-funded by MRC and Roche, UCB, Eli Lilly and Novartis); Joint PhD funding from EPSRC and AZ, and grant funding from Vistagen Therapeutics. He has also unrestricted educational grant support for the UK Pharmacogenetics and Stratified Medicine Network from Bristol-Myers Squibb and UCB. He has developed an HLA genotyping panel with MC Diagnostics, but does not benefit financially from this. MP is part of the IMI Consortium ARDAT (www.ardat.org). SQ has received fees from GSK, AZ, Sanofi, Teva, Novartis, and Chiesi. SJ-HV has received grants from SysPharmPediA EraNet. VdP has received fees from AZ and GSK. VP has received fees from Sanofi, AZ, Chiesi, MSD, and Boehringer Ingelheim, grant support from MSD, Chiesi Institutional, and Menarini. EvM has received grants from the German Federal Ministry of Education and Research and the Bavarian State Ministry of Health and Care, royalties/licenses from Elsevier GmbH, Georg Thieme Verlag, Springer-Verlag GmbH and Elsevier Ltd. EvM has recieved fees from the Chinese University of Hongkong, European Commission, HiPP GmbH & Co KG, AZ, Imperial College London, Massachusetts Medical Society, Springer-Verlag GmbH, Elsevier Ltd., Böhringer Ingelheim International GmbH, European Respiratory Society (ERS), Universiteit Utrecht, Faculteit Diergeneeskunde, Universität Salzburg, Springer Medizin Verlag GmbH, Japanese Society of Pediatric Allergy and Clinical Immunology (JSPACI), Klinikum Rechts der Isar, University of Colorado, Paul-Martini-Stiftung, Astra Zeneca, Imperial College London, Children´s Hospital Research Institute of Manitoba, Kompetenzzentrum für Ernährung (Kern), OM Pharma S.A., Swedish Pediatric Society for Allergy and Lung Medicine, Chinese College of Allergy and Asthma (CCAA), Verein zur Förderung der Pneumologie am Krankenhaus Großhansdorf e.V., Pneumologie Developpement, Mondial

Congress & Events GmbH & Co. KG, American Academy of Allergy, Asthma & Immunology, Imperial College London, Margaux Orange, Volkswagen Stiftung, Böhringer Ingelheim International GmbH, European Respiratory Society (ERS), Universiteit Utrecht, Faculteit Diergeneeskunde, Österreichische Gesellschaft f. Allergologie u. Immunologie, Massachusetts Medical Society, OM Pharma S. A., Hanson Wade Ltd., iKOMM GmbH, DSI Dansk Borneastma Center, American Thoracic Society, HiPP GmbH & Co KG, Universiteit Utrecht, Faculteit Bètawetenschappen. EvM has patents No. PCT/EP2019/085016, EP21189353.2. 2021. and PCT/US2021/016918. 2021. pending, royalties paid to ProtectImmun for patent EP2361632 and patents EP1411977, EP1637147, and EP 1964570 licensed to ProtectImmun. EvM is a member of the EXPANSE Scientific Advisory Board, Member of the BEAMS External Scientific Advisory Board (ESAB), Member of the Editorial Board of "The Journal of Allergy and Clinical Immunology: In Practice", Member of the Scientific Advisory Board of the Children's Respiratory and Environmental Workgroup (CREW), Member of the International Scientific & Societal Advisory Board (ISSAB) of Utrecht Life Sciences (ULS), University of Utrecht, Member of External Review Panel of the Faculty of Veterinary Science, University of Utrecht, Member of the Selection Committee for the Gottfried Wilhelm Leibniz Programme (DFG), Member of the International Advisory Board of Asthma UK Centre for Applied Research (AUKCAR), Member of the International Advisory Board of "The Lancet Respiratory Medicine", Member of the Scientific Advisory Board of the CHILD (Canadian Healthy Infant Longitudinal Development) study, McMaster University, Hamilton, Canada, Asthma UK Centre for Applied Research and the Pediatric Scientific Advisory Board Iceland. The other authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

EH-L: Conceptualization (supporting), Data curation (equal), Formal analysis (lead), Methodology (equal), Writing original draft (lead), Writing review and editing (equal); **VEO**: Resources (equal), Writing-review and editing (supporting); funding acquisition (lead); **EJA**: Resources (equal), Data curation (equal), Formal analysis (supporting), Writing-review and editing (supporting); **YYS**: Resources (equal), Investigation (supporting), Data curation (equal), Formal analysis (supporting), Writing-review and editing (supporting); Resources (equal), Investigation (supporting), Data curation (equal), Formal analysis (equal), Writing-review and editing (supporting); **RG**: Resources (equal), Data curation (equal), Formal analysis (supporting), Writing-review and editing (supporting); **EDR**: Data curation (equal), Formal analysis (supporting), Writing-review and editing (supporting); **NT**: Data curation (equal), Formal analysis (supporting), Writing-review and editing (supporting); **EEV**: Resources (equal), Investigation (supporting), Data curation (equal), Formal analysis (supporting), Writing-review and editing (supporting); **NH-P**: Investigation (supporting), Data curation (equal), Formal analysis (supporting), Writing-review and editing (supporting); **JP-G**: Investigation (supporting), Formal analysis (supporting), Writing review and editing (equal); **EM-G**: Investigation (supporting), Writing review and editing (supporting); **FL-D**: Resources (equal), Conceptualization (supporting), Funding acquisition (supporting); **SH**: Resources (equal), Data

curation (equal), Writing-review and editing (supporting); **PB**: Resources (equal), Data curation (equal), Writing-review and editing (supporting); **AJ**: Data curation (equal), Formal analysis (supporting), Writing-review and editing (supporting); **QY**: Formal analysis (supporting), **EF**: Data curation (equal), Formal analysis (supporting), Writing-review and editing (supporting); **SJV**: Resources (equal), Data curation (equal), Writing-review and editing (supporting); **RL**: Data curation (equal), Writing-review and editing (supporting); **AE-O**: Formal analysis (supporting), Writing-review and editing (supporting); **MG**: Resources (equal), Investigation (supporting), Data curation (equal), Formal analysis (supporting), Writing-review and editing (supporting); **CE**: Investigation (supporting), Writing-review and editing (supporting); **RG-P**: Resources (equal), Investigation (supporting), Writing-review and editing (supporting); **JMH-P**: Resources (equal), Investigation (supporting), Writing-review and editing (supporting); **PP-G**: Resources (equal), Investigation (supporting), Writing-review and editing (supporting); **OS**: Resources (equal), Investigation (supporting), Writing-review and editing (supporting); **PC**: Resources (equal), Investigation (supporting), Writing-review and editing (supporting); **GAH**: Investigation (supporting), Writing-review and editing (supporting); **AM**: Data curation (supporting), Writing review and editing (supporting); **TB**; Investigation (supporting), Writing-review and editing (supporting); **KFR**: Investigation (supporting), Writing-review and editing (supporting); **GH**: Investigation (supporting), Writing-review and editing (supporting); **MVK**: Investigation (supporting), Writing-review and editing (supporting); **RR**: Formal analysis (supporting), Writing-review and editing (supporting); **MJC**: Investigation (supporting), Writing review and editing (supporting); **FJG-B**: Investigation (supporting), Writing review and editing (supporting); **JMO**: Investigation (supporting), Writing review and editing (supporting); **VP**: Investigation (supporting), Writing review and editing (supporting); **SQ**: Investigation (supporting), Writing review and editing (supporting); **GC**: Investigation (supporting), Writing review and editing (supporting); **MC**: Investigation (supporting), Writing review and editing (supporting); **VdP**: Resources (supporting), Investigation (supporting), Writing-review and editing (supporting); **JRR-S**: Investigation (supporting), Writing-review and editing (supporting); **JK-M**: Resources (equal), Investigation (supporting), **JV**: Conceptualization (supporting), Resources (supporting), Writing-review & editing (equal); **UP**: Resources (equal), Writing-review & editing (supporting); **CF**: Resources (equal), Data curation (equal), Formal analysis (supporting), Writing-review and editing (supporting); **MK**: Resources (equal), Data curation (equal), Writing-review and editing (supporting); **SM**: Resources (equal), Data curation (equal), Writing-review and editing (supporting); **MP**: Resources (equal), Data curation (equal), Writing-review and editing (supporting); **DBH**: Resources (equal), Data curation (equal), Writing-review and editing (supporting); **EvM**: Resources (equal), Writing-review & editing (supporting); **CNP**: Resources (equal), Writing-review & editing (supporting); **ST**: Resources (equal), Writing-review & editing (supporting); **A-HMZ**: Resources (supporting), Writing-review and editing (supporting); **EM**: Resources (equal), Writing-review and editing (supporting); **JC**: Resources (equal), Writing-review and editing (supporting); **GB**: Resources

(equal), Writing-review and editing (supporting); **FTC**: Resources (equal), Writing-review and editing (supporting); **EB**: Resources (equal), Writing-review & editing (supporting); **DM:** Resources (equal), Writing-review & editing (supporting); **EGB**: Resources (equal), Funding acquisition (lead), Supervision (supporting), Writing-review & editing (supporting); **MP-Y**: Resources (equal), Conceptualization (lead), Supervision (lead), Funding acquisition (lead); Methodology (equal), Writing original draft (supporting), Writing review and editing (lead).

ABSTRACT

Background: Asthma exacerbations are a serious public health concern due to high health care resource utilization, work/school productivity loss, impact on quality of life, and risk of mortality. The genetic basis of asthma exacerbations has been studied in several populations, but no prior study has performed a multi-ancestry meta-analysis of genome-wide association studies (meta-GWAS) for this trait. We aimed to identify common genetic loci associated with asthma exacerbations across diverse populations and to assess their functional role in regulating DNA methylation and gene expression.

Methods: A meta-GWAS of asthma exacerbations in 4,989 Europeans, 2,181 Hispanics/Latinos, 1,250 Singaporean Chinese, and 972 African Americans analyzed 9.6 million genetic variants. Suggestively associated variants ($p \leq 5x10^{-5}$) were assessed for replication in 36,477 European and 1,078 non-European asthma patients. Functional effects on DNA methylation were assessed in 595 Hispanic/Latino and African American asthma patients and in publicly available databases. The effect on gene expression was evaluated *in silico*.

Results: 126 independent variants were suggestively associated with asthma exacerbations in the discovery phase. Two variants independently replicated: rs12091010 located at vascular cell adhesion molecule-1/exostosin like glycosyltransferase-2 (*VCAM1/EXTL2*) (discovery: odds ratio ($OR_{T\ allele}$) =0.82, p=9.05x10^{-6} and replication: $OR_{T\ allele}$=0.89, p=5.35x10^{-3}) and rs943126 from pantothenate kinase1 (*PANK1*) (discovery: $OR_{C\ allele}$=0.85, p=3.10x10^{-5} and replication: $OR_{C\ allele}$=0.89, p=1.30x10^{-2}). Both variants regulate gene expression of genes where they locate and DNA methylation levels of nearby genes in whole blood.

Conclusions: This multi-ancestry study revealed novel regulatory loci for asthma exacerbations located in genomic regions participating in inflammation and host defense.

KEYWORDS

Asthma exacerbations, *EXTL2,* GWAS, single nucleotide polymorphism, *PANK1*

ABBREVIATIONS

1KGP: 1000 Genomes Project

CDK6: Cyclin-dependent kinase 6

CI: Confidence interval

GAG: Glycosaminoglycan

GTEx: Genotype-Tissue Expression

GWAS: Genome-wide association study

LPS: Lipopolysaccharide

MAF: Minor allele frequency

meQTL: Methylation quantitative trait loci

OR: Odds ratio

PPAR-α: Peroxisome proliferator-activating receptor α

RR: Relative risk

SNP: Single nucleotide polymorphism

TLR4: Toll-like receptor 4

TNFα: Tumor necrosis factor α

INTRODUCTION

Asthma is a common chronic inflammatory airway disorder affecting over 300 million people worldwide. The disparities in asthma prevalence across populations reflect a complex interplay between environmental exposures (i.e., air pollution and viral infections), behavioral and socioeconomic factors (i.e., treatment adherence and healthcare access), and genetic ancestry, which is a complex trait measured by background whole-genome variation that tracks with geographic and historical factors as well as the aforementioned factors influencing asthma prevalence (1,2).

Asthma exacerbations are defined as worsening of respiratory symptoms requiring hospitalization, unscheduled/emergency asthma care, and/or use of systemic corticosteroids (3). Prevention of asthma exacerbations is a major public health priority due to their associated consequences on health (i.e., decreased quality of life, accelerated decline in lung function, or mortality), school attendance, work productivity, and healthcare costs (1,4,5). To date, the best predictor of future exacerbations is the occurrence of one in the previous year (6). Thus, identifying potential biomarkers to guide the reduction and prevention of exacerbations is a priority for therapeutics development and for precision medicine of asthma.

With the advent of high-throughput sequencing and genotyping technologies, the study of the genetic contributions to asthma exacerbations has shifted from hypothesis-driven, limited candidate-gene strategies to genome-wide association studies (GWAS) (7). Pharmacogenomics studies of asthma exacerbations as an outcome of treatment response have identified five suggestive associations for asthma exacerbations despite inhaled corticosteroids (*CMTR1* (8), *APOBEC3B-APOBEC3C* (9), and *CACNA2D3-WNT5A* (10)), or long-acting beta2-agonists (*TBX3* and *EPHA7*) (11). Beyond pharmacogenomics, other studies have focused on asthma exacerbations independently of treatment. In European-descent populations, *CDHR3, CTNNA3*, and *HLA-DQB1* have been associated with severe asthma exacerbations (12,13). More recently, the representation of ethnically diverse populations has increased in GWAS of asthma exacerbations. A meta-analysis of GWAS in Hispanic/Latino children identified a single nucleotide polymorphism (SNP) at *FLJ22447* that modulated *KCNJ2-AS1* expression in nasal epithelium through DNA methylation (14). In Hispanic/Latinos and African Americans, a genome-wide significant locus for asthma with exacerbations regulated *LINC01913* lung gene expression and DNA methylation levels of the *PKDCC* gene in whole blood (15). However, none of those studies has approached the search for genetic determinants of asthma exacerbations independently of treatment from a multi-ancestry framework.

To improve our understanding on genetic and biological mechanisms of asthma exacerbations across multiple populations, we conducted the first multi-ethnic meta-analysis of GWAS of asthma exacerbations and attempted to validate previous associations. Then, we conducted *in silico* and

in vivo downstream analyses to assess the potential functional effects of the associated SNPs over DNA methylation and gene expression.

METHODS

Study design and study populations

We performed a two-stage study to identify genetic variants associated with asthma exacerbations, defined as a binary variable based on the presence of emergency care, hospitalizations, or administration of systemic corticosteroids because of asthma. We also considered a definition of moderate exacerbations (3), comprising unscheduled general practitioner or pulmonary specialist visits and school absence since no information on the former variables was available for some studies. A period of 6 to 24 months was considered depending on the data available for each study (**Table S1-S2**). In the discovery phase, we performed a multi-ethnic meta-analysis of GWAS of asthma exacerbations in 9,392 patients with asthma from 12 studies, including 4,989 European-descents from nine studies, 2,181 Hispanics/Latinos, 1,250 Singaporean Chinese, and 972 African Americans. We attempted to replicate the findings from the discovery phase in a total of 37,555 participants with asthma, including 36,477 Europeans from seven studies, 877 Latinos from two studies, and 201 Filipinos from one study (**Table S2**). A detailed description of each study is available in the **Supporting Information**. All studies included were approved by their respective Institutional review boards, and written informed consent was provided by participants or their parents/caregivers. All methods followed the Declaration of Helsinki guidelines.

Assessment of genetic ancestry was performed using principal component analysis. The Haplotype Reference Consortium (r1.1 2016) (16) was used as the reference imputation panel for most studies, except for Avon Longitudinal Study of Parents and Children (ALSPAC) and Singapore Cross Sectional Genetic Epidemiology Study (SCSGES), which used the phase 3 of the 1000 Genomes Project (1KGP) (17). Genotyping and imputation procedures for the discovery and replication studies are detailed in the **Supporting Information** and **Tables S1-S2.**

Association analysis

Association between genetic variants and asthma exacerbations was tested using logistic regression models including age, sex, and principal components from the genotype matrix (if needed to correct for population stratification) (**Table S1**)**.** Analyses were conducted separately for each study using PLINK 2.0 (18), EPACTS 3.2.6 (19) or rvtests 2.1.0 (20). Results were filtered with the EasyQC software (21) to retain variants with a minor allele frequency (MAF)≥1% and imputation quality $R^2 \geq 0.3$, absolute value of the beta coefficient <10, standard error of the beta included in the interval [0,10], and minor allele cut-off ≥6.

In the discovery phase, genetic variants that were available in at least two ethnic-specific studies were meta-analyzed with METASOFT (22), using fixed-effects or random-effects models based on the heterogeneity among studies (measured by the Cochran's Q test *p*-value). Ethnic-specific results were then combined in a multi-ethnic meta-analysis. Independent variants ($r^2 \leq 0.8$) with suggestive association at $p \leq 5x10^{-5}$ within 1 Megabase were identified with GCTA-COJO v1.93.2 (23) using the 1KGP reference (17). These variants were evaluated in the replication stage, following the same procedures as in the discovery phase. Evidence of replication was considered if the variants showed consistent direction of effects with the discovery stage at $p \leq 0.05$.

Assessment of shared genetic basis of asthma exacerbations with other traits

To identify groups of genes previously associated with other traits, we used a Gene-Set Enrichment Analysis (GSEA), as implemented in FUMA GWAS (24) via the *GENE2FUNC* algorithm, and queried the GWAS catalog (25). SNPs with $p \leq 1x10^{-4}$ in the discovery phase of the meta-analysis of GWAS were mapped to the closest gene using the UCSC Table Browser tool (26). A false discovery rate (FDR) of 5% was used to declare significance.

To estimate the pairwise genome-wide genetic correlations (R_g) between asthma exacerbations and other traits, we compared our findings with publicly-available GWAS summary statistics via LD score regression using LDHub (27). Since most of the GWAS have been conducted in European populations, the analysis was restricted to predominantly European-descent individuals to maximize the statistical power. A Bonferroni-corrected significance threshold of $p<0.05/711$ traits=$6.48x10^{-5}$ was applied.

Sensitivity analysis

In order to assess the robustness of the genetic associations, we conducted sensitivity analyses for the time-dependent probability occurrence of exacerbations, and the effect of Body Mass Index (BMI), obesity and asthma severity (defined as detailed in the **Supporting Information)**. Studies from the discovery stage that had covariate data available were considered.

Methylation profiling and quality control

Whole blood DNA methylation from Hispanics/Latinos and African Americans was profiled using the Infinium HumanMethylation450 BeadChip or the Infinium Methylation EPIC BeadChip arrays. Briefly, low-quality probes and samples, outliers of DNA methylation, and samples with sex mismatch or mixed genotype distributions on the control SNP probes were excluded. Standard background correction, dye-bias correction, inter-array normalization, and probe-type bias adjustment were performed, and beta values were transformed to M-values for better statistical performance. Quality control is detailed in the **Supporting Information**.

Functional assessment of associated SNPs

DNA methylation quantitative trait loci (meQTL) analyses were conducted using fastQTL (28) for CpG sites within 1 Mb of SNPs with MAF≥0.01 in at least 10 samples, separately in 139 Mexican

Americans and 241 Puerto Ricans from Genes-Environments & Admixture in Latino Americans (GALA II) and 215 African Americans from the Study of African Americans, Asthma, Genes & Environments (SAGE) studies. Population-specific results were then meta-analyzed with METASOFT (22). Linear regression models were corrected for asthma exacerbations status, age, sex, genetic ancestry, ReFACTor components as a proxy of cell heterogeneity, and methylation batch (when appropriate). *In silico* evidence of functional effects of variants on gene expression and DNA methylation was assessed using QTLbase (29), Genotype-Tissue Expression (GTEx) v8 Portal (30), PhenoScanner v2 (31) and eFORGE-TF (32). Long-distance chromatin interactions were determined using the ChiCP tool (33).

Validation of previous associations

A literature search for all studies reporting genetic loci significantly associated with asthma exacerbations was conducted, as described in the **Supporting Information**. Association results in the discovery stage were extracted and significance threshold was defined as p = 0.05/ number of tested SNPs to adjust for multiple testing.

RESULTS

Characteristics of the patients

In the discovery phase, we analyzed 2,781 exacerbators and 6,611 non-exacerbators; 53.1% were predominantly Europeans, 23.2% Hispanics/Latinos, 13.3% Singaporean Chinese, and 10.3% African Americans. The percentage of exacerbators ranged from 9.1% to 65.2% in Europeans, and reached 58.8% in Hispanics/Latinos, 46.1% in African Americans, and 3.4% in Singaporeans. The replication phase included 37,555 individuals with asthma (3,030 exacerbators and 34,525 non-exacerbators) where most participants were of European-descent (97.1%), followed by Latinos (2.3%) and Filipinos (0.5%). The percentage of exacerbators ranged from 4.8% to 65.2% in Europeans, reached approximately 43% in Latinos, and 1.3% in Filipinos (**Table S1-S2**). Regarding sex, 51.7% and 42.9% of participants were male in the discovery and replication phases, respectively.

Discovery phase

The Quantile-quantile plots did not show major genomic inflation due to population stratification in each individual study (**Figure S1**), the combined results from individuals of European descent (**Figure S2**), or the multi-ancestry meta-analysis (**Figure S3**). In the multi-ancestry meta-analysis of 9,634,748 variants, 447 SNPs exhibited suggestive association (**Table S3**). Chromosome 5p14.3 harbored a genome-wide significant (**Figure 1, Figure S4**) intronic SNP rs6888198 within the cadherin-12 (*CDH12*) gene as the most significant association (odds ratio [OR] for C allele: 1.37, 95% confidence interval [CI]: 1.23-1.54, p=1.95x10^{-8}).

Replication phase

Fifteen of the 126 independent variants identified in the discovery phase were not available for replication since they were mostly present in African Americans and Hispanics/Latinos (**Table S3**). Two of the 106 variants present in more than one ethnic group were consistently associated with asthma exacerbations (**Table 1**): rs12091010 [*VCAM1/EXTL2,* OR for T allele: 0.89 (0.82–0.97), p=5.35x10^{-3}] (**Figure 2**) and rs943126 [*PANK1*, OR for C allele: 0.92 (0.86–0.98), p=1.30x10^{-2}] (**Figure 3**). In the meta-analysis across both phases, these variants reached an association *p*-value of 4.23x10^{-7} and 4.93x10^{-6}, respectively. From five variants that were present only in non-Europeans in the replication stage, none exhibited $p<0.05$ in any other population group (**Table S4**). Even though rs6888198 reached genome-wide significance in the discovery and showed consistent effects among Europeans in the replication phase, this SNP had opposite effects in Latinos and Filipinos, which resulted in the lack of replication in the multi-ancestry replication phase (**Table 1, Figure S5**).

Gene-set enrichment and genome-wide genetic correlation analysis

Enrichment analysis of associations from the multi- ancestry discovery GWAS including 959 SNPs associated with asthma exacerbations at $p\leq 1\text{x}10^{-4}$ revealed significant enrichment in several traits, including treatment response (min p=2.77x10^{-6}), neurological conditions (min p=4.62x10^{-5}), obesity (min p=6.52x10^{-5}), or waist-to-hip ratio (min p=1.88x10^{-7}) (**Table S5**).

A total of 16 traits exhibited genetic correlation with asthma exacerbations at $p<0.05$ (**Table S6**), including wheeze or whistling in the last year (R_g=0.47, p=1.01x10^{-2}), emphysema/chronic bronchitis (R_g=0.55, p=3.89x10^{-2}), asthma (R_g=0.32, p=3.99x10^{-2}), and BMI (R_g=0.19, p=4.76x10^{-2}). However, the associations did not remain significant after Bonferroni correction.

Sensitivity analysis

To assess the robustness of associations that replicated across stages to the time-dependent probability of occurrence of exacerbations, stratified analyses were performed in European-descents from the discovery stage that reported exacerbations for 6 *vs*. 12 months. Consistent effects per period were observed across periods (**Table 3**).

Since the post-GWAS analyses revealed significant enrichment/correlation at p<0.05 with fat mass/distribution, the association of rs12091010 and rs943126 after additional adjustment by BMI/obesity was examined in individuals from the discovery phase with BMI data available. Moreover, the effect of asthma severity alone or combined with BMI/obesity on the genetic association exacerbations was evaluated. The effects sizes of the genetic association after additional adjustment by these variables remained consistent with the effects reported in the discovery stage (**Table S7**).

Functional exploration of variants associated with asthma exacerbations

We next assessed for association DNA methylation in whole blood at 525, 538, and 99 CpG sites with rs12091010, rs943126, and rs6888198, respectively. A total of 7 and 1 SNP-CpG pairs for rs943126 and rs12091010 exhibited Storey $q<0.05$, respectively (**Table 2, Table S8**). Two of these replicated consistently in Europeans for rs943126 (cg25770176 and cg00475140). *In silico* analyses, revealed 10 SNP-CpGs pairs, 3 of which showed consistent effects in Hispanics/Latinos and African Americans at Storey $q<0.05$ (**Tables S9-S10**) including the previous two pairs and rs943126-cg03948048. The 8 significant CpG sites in minority children showed significant enrichment (q<0.001) in transcription factor (TF) motifs in lung (**Table S11**). These TFs where enriched (q<0.05) in biological processes including regulation of biosynthetic processes, transcription by RNA polymerase II, hormone metabolic process or response to endogenous stimulus, and molecular functions including binding to a regulatory nucleic acid region or protein homo/heterodimerization activity (**Table S12**). Besides, the T allele of rs12091010 was associated with decreased *EXTL2* expression in whole blood from Europeans, according to PhenoScanner (31). The C allele of rs943126 was associated with increased expression of *PANK1* in whole blood from Europeans (**Table S13**). Both variants showed evidence of long-range chromatin interaction with several genes in lymphoblastoid cells, including *VCAM1* and *EXTL2* for rs12091010 and *PANK1* for rs943126 (**Table S14**).

Validation of previous associations

We next examined 47 previous genetic loci for asthma exacerbations (9,12–14,34–36) and moderate-to-severe asthma (37) for association with asthma exacerbations in the discovery phase. A total of 5 variants had $p<0.05$ in Europeans, 3 in Hispanics/Latinos, 5 in African Americans, and 1 in Singaporean Chinese (**Table S15**). These were in loci previously associated with asthma exacerbations (*GSDMB, RAD50, HLA-DQB1, ADAM33, VDR*, and *CDHR3*) or moderate-to-severe asthma (*IKZF3, TSLP, MUC5AC, C11orf30, SMAD3*, and *WDR36*). However, none of the SNPs exceeded the stringent Bonferroni-corrected threshold for significance ($p=0.05/47=1.06\times10^{-3}$).

DISCUSSION

To our knowledge, this is the first multi-ancestry meta-analysis of GWAS of asthma exacerbations including European, Hispanic/Latino, Asian, and African American patients with asthma. In our combined analysis of 46,947 individuals with asthma, two regulatory SNPs were significantly and consistently associated with asthma exacerbations in most of the studies included in the discovery and replication phases, independently of the type of exacerbation and the time period for which the exacerbation status was assessed. The SNP rs120910109 was located in the intergenic region of the *VCAM1/EXTL2* genes whereas rs943126 was harbored within an intron 1 of *PANK1.*

VCAM1 encodes a surface protein predominantly expressed in endothelial cells that modulates leukocyte adhesion and trans-endothelial migration in response to pro-inflammatory cytokines, and lipopolysaccharide (LPS) among other factors (38,39). VCAM1 is involved in cancer progression and several immunological disorders, including asthma (38). In the ovalbumin mice model, anti-VCAM1 reduced airway hyperresponsiveness and eosinophilic inflammation (40). On the other hand, *EXTL2* encodes an enzyme that controls glycosaminoglycan (GAG) biosynthesis via transference of N-acetylgalactosamine and N-acetylglucosamine to the glycosaminoglycan-protein linkage region (41). Decreased *EXTL2* causes an over-accumulation of GAGs(42) that can promote inflammation in injured areas (43,44). Moreover, in bone marrow-derived macrophages from *EXTL2*$^{-/-}$ mice, there is overproduction of key molecules involved in inflammation and extracellular matrix remodeling, including tumor necrosis factor α (TNFα) and several matrix metalloproteinases (43). In a scenario of overaccumulation of GAGs under the loss of *EXTL2* in macrophages, GAGs act as inflammatory mediators with strong Toll-like receptor 4 (TLR4) agonist capacity (44). Interestingly, genetic variation both *VCAM1* and *EXTL2* is associated with blood cell counts, and multiple sclerosis, according to the GWAS catalog (25).

PANK1 catalyzes coenzyme A biosynthesis, regulated by the transcription factor peroxisome proliferator-activating receptor α (PPAR-α) (45), a key anti-inflammatory factor in asthma (46). A decrease in PPAR-α expression is accompanied by a decrease in the expression of PANK1 and miR-107, which is encoded within the intron 5 of *PANK1*. TLR4 can also downregulate miR-107. In turn, this leads to a higher cyclin-dependent kinase 6 (CDK6) expression and subsequently increases the adhesion of macrophages in response to LPS (45). Bioproducts from bacterial infections, such as LPS, can trigger an inflammatory response and increase airway hyperresponsiveness and risk of asthma exacerbations (47,48). Moreover, p53 can regulate cell cycle progression via upregulation of PANK1 after DNA damage (49) and metabolism (50).

To prioritize gene targets, we assessed the functional capacity of relevant SNPs (51). Both rs12091010 and rs943126 exhibited an association with DNA methylation at several nearby CpG sites in whole blood from African Americans and Hispanics/Latinos with asthma. Additionally, the SNPs rs12091010 and rs943126 were associated with *EXTL2* and *PANK1* gene expression in whole blood from Europeans. Specifically, the T allele of rs12091010, located at 6 kb downstream of the 3' UTR of *VCAM1* and 150 kb upstream of the transcription start site of *EXTL2,* was associated with lower odds of having asthma exacerbations and decreased *EXTL2* expression (31) The T allele is more common among Latinos/Admixed Americans, followed by Europeans, Africans, and East Asians (**Figure S6**). The T allele of rs943126 at *PANK1,* which is less common among Europeans than the rest of populations (**Figure S7**), was associated with a higher risk of asthma exacerbations in the combined analysis of the discovery and replication phases and with decreased gene expression of *PANK1* in whole blood from Europeans. However, these eQTL effects were not validated in the GTEx data (30).

In the discovery phase, the intronic SNP rs6888198 (*CDH12)* locus exceeded the genome-wide significance threshold but showed non-significant association (p>0.05) in the replication phase, although the effect of the association was consistent across study phases. Furthermore, rs6888198 showed variable MAF among populations, with the largest MAF among Africans and Latinos (**Figure S8**). *CDH12* has been associated with angiogenesis and progression of several types of cancers (52–54). Specifically, in colorectal cancer, it has been suggested that *CDH12* increases cancer cell migration by promoting epithelial-mesenchymal transition via activation of the Snail transcription factor pathway. *CDH12* expression is positively modulated by the chemotactic factor *CCL2* (53,54), whose levels increases in blood and airway smooth muscle from asthma patients compared to healthy controls (55).

We also attempted to assess previously associated loci for asthma exacerbations or moderate-to-severe asthma for association with asthma exacerbations in multiple ethnic groups. Two variants were consistently associated with the risk of asthma exacerbations in more than one ethnic group: rs72743461 (*SMAD3*) and rs9275503 (*HLA-DQB1*) in both Europeans and Hispanics/Latinos. Only the *TSLP* locus rs1837253 was associated at p<0.05 with asthma exacerbations in Singaporean Chinese. Of note, none of the previous findings was initially described in Asian or African populations, which highlights the need to increase ethnic diversity in genomic studies of asthma exacerbations.

Our study has several limitations. First, the *VCAM1/EXTL2* and *PANK1* loci did not reach genome-wide significance in the combined analysis from all studies. Second, the history of asthma exacerbations was based on self-reported retrospective questionnaires in all cohorts but COMPASS, a randomized, prospective clinical trial. Third, to bring together large sample sizes necessary to map susceptibility variants, we considered studies where asthma exacerbations were reported for the previous 6 to 24 months, which may have introduced some heterogeneity in the phenotype. Despite these limitations, our findings exhibited consistent effects for the *VCAM1/EXTL2* and *PANK1* loci independent of the time period assessed. Future studies should untangle the role of this loci in the time-to-first exacerbation, the annual number of exacerbations, or the temporal distance among events and further explore the biological function of these genes. On the other hand, we acknowledge several study strengths. Firstly, we leveraged clinical and genetic data from 46,947 asthma patients from different ethnicities from 18 independent studies. Our study had statistical power ≥80% to detect associations with MAF> 17% and relative risk (RR)>1.20 in the discovery stage and for variants with MAF≥1%, and was powered at 80% to detect associations with larger effect sizes (RR≥1.85). Second, we identified novel, biologically plausible genetic factors of asthma exacerbations demonstrated by transcriptomics and epigenomics studies and evidence for prior literature. Third, we evaluated previous genetic signals from asthma exacerbations in populations from several ancestries.

We demonstrate for the first time the power of multi-ancestry GWAS to successfully identify loci for asthma exacerbations with consistent genetic effects across individuals from varying ancestral backgrounds. We also demonstrated that these loci are biologically functional and regulate RNA expression and adjacent CpG site DNA methylation as meQTL in whole blood cells. Our findings highlight *VCAM1, EXTL2* and *PANK1* as functional loci for asthma exacerbations applicable to people across different ancestral backgrounds warranting future investigation of these novel genomic mechanisms underlying asthma exacerbations.

DATA AVAILABILITY

All data necessary to evaluate the conclusions of this manuscript are reported in the main text and/or the Supporting Information. Genome-wide genotyping data for GALA II and SAGE are available at the database of Genotypes and Phenotypes (dbGaP) (Study Accession phs001274.v2.p1 and phs000092.v1.p1, respectively). The summary statistics of the multi-ancestry discovery phase are available at the Zenodo repository: 10.5281/zenodo.5513443.

IMPACT STATEMENT

A large multi-ancestry meta-analysis of GWAS of asthma exacerbations revealed two novel susceptibility loci located close to *PANK1* and at the intergenic region of *VCAM1* and *EXTL2*. These loci decreased *PANK1* and *EXTL2* gene expression in whole blood, respectively. Both genetic variants were associated with DNA methylation levels at CpG sites nearby. Our results identified two gene targets for asthma exacerbations that should be further explored to assess their specific role in asthma.

REFERENCES

1. Global Initiative for Asthma. Global strategy for asthma management and prevention. 2021.http://ginasthma.org/ (accessed 21 Sep2020).
2. Hernandez-Pacheco N, Flores C, Oh SS, Burchard EG, Pino-Yanes M. What Ancestry Can Tell Us About the Genetic Origins of Inter-Ethnic Differences in Asthma Expression. *Curr Allergy Asthma Rep* 2016;**16**:53.
3. Reddel HK, Taylor DR, Bateman ED, Boulet LP, Boushey H a., Busse WW et al. Asthma control and exacerbations - Standardizing endpoints for clinical asthma trials and clinical practice. *Am J Respir Crit Care Med* 2009;**180**:59–99.
4. Calhoun WJ, Haselkorn T, Miller DP, Omachi TA. Asthma exacerbations and lung function in patients with severe or difficult-to-treat asthma. *J Allergy Clin Immunol* 2015;**136**:1125-1127.e4.
5. Chipps BE, Haselkorn T, Rosén K, Mink DR, Trzaskoma BL, Luskin AT. Asthma Exacerbations and Triggers in Children in TENOR: Impact on Quality of Life. *J Allergy Clin Immunol Pract* 2018;**6**:169-176.e2.
6. Puranik S, Forno E, Bush A, Celedón JC. Predicting severe asthma exacerbations in children. Am. J. Respir. Crit. Care Med. 2017;**195**:854–859.
7. Herrera-Luis E, Hernandez-Pacheco N, Vijverberg SJ, Flores C, Pino-Yanes M. Role of genomics in asthma exacerbations. *Curr Opin Pulm Med* 2019;**25**:101–112.
8. Dahlin A, Denny J, Roden DM, Brilliant MH, Ingram C, Kitchner TE et al. CMTR1 is associated with increased asthma exacerbations in patients taking inhaled corticosteroids. *Immunity, Inflamm Dis* 2015;**3**:350–359.
9. Hernandez-Pacheco N, Farzan N, Francis B, Karimi L, Repnik K, Vijverberg SJ et al. Genome-wide association study of inhaled corticosteroid response in admixed children with asthma. *Clin Exp Allergy* 2019;**49**:789–798.
10. Hernandez-Pacheco N, Vijverberg SJ, Herrera-Luis E, Li J, Sio YY, Granell R et al. Genome-wide association study of asthma exacerbations despite inhaled corticosteroid use. *Eur Respir J* 2021;**57**.
11. Slob EMA, Richards LB, Vijverberg SJH, Longo C, Koppelman GH, Pijnenburg MWH et al. Genome-wide association studies of exacerbations in children using long-acting beta2-agonists. *Pediatr Allergy Immunol* Published Online First: 11 March 2021. doi:10.1111/pai.13494
12. Yan Q, Forno E, Herrera-Luis E, Pino-Yanes M, Yang G, Oh S et al. A genome-wide association study of asthma hospitalizations in adults. *J Allergy Clin Immunol* 2021;**147**:933–940.
13. Herrera-Luis E, Hernandez-Pacheco N, Vijverberg SJ, Flores C, Pino-Yanes M. Role of

genomics in asthma exacerbations. *Curr Opin Pulm Med* 2019;**25**:101–112.

14. Yan Q, Forno E, Herrera-Luis E, Pino-Yanes M, Qi C, Rios R et al. A genome-wide association study of severe asthma exacerbations in Latino children and adolescents. *Eur Respir J* 2021;**57**:2002693.
15. Herrera-Luis E, Espuela-Ortiz A, Lorenzo-Diaz F, Keys KL, Mak ACY, Eng C et al. Genome-wide association study reveals a novel locus for asthma with severe exacerbations in diverse populations. *Pediatr Allergy Immunol* 2021;**32**:106–115.
16. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 2016;**48**:1279–1283.
17. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM et al. A global reference for human genetic variation. *Nature* 2015;**526**:68–74.
18. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* 2015;**4**.
19. Kang HM. EPACTS (Efficient and Parallelizable Association Container Toolbox). Accessed March 24, 2019. https://genome.sph.umich.edu/wiki/EPACTS. 2016.
20. Zhan X, Hu Y, Li B, Abecasis GR, Liu DJ. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics* 2016;**32**:1423–1426.
21. Winkler TW, Day FR, Croteau-Chonka DC, Wood AR, Locke AE, Mägi R et al. Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* 2014;**9**:1192–1212.
22. Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am J Hum Genet* 2011;**88**:586–598.
23. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011;**88**:76–82.
24. Watanabe K, Taskesen E, Van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 2017;**8**:1826.
25. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 2019;**47**:D1005–D1012.
26. Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D et al. The UCSC Table Browser data retrieval tool. *Nucleic Acids Res* 2004;**32**:D493-6.
27. Zheng J, Erzurumluoglu AM, Elsworth BL, Kemp JP, Howe L, Haycock PC et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* 2017;**33**:272–279.

28. Ongen H, Buil A, Brown AA, Dermitzakis ET, Delaneau O. Fast and efficient QTL mapper for thousands of molecular phenotypes. *Bioinformatics* 2016;**32**:1479–1485.

29. Zheng Z, Huang D, Wang J, Zhao K, Zhou Y, Guo Z et al. QTLbase: an integrative resource for quantitative trait loci across multiple human molecular phenotypes. *Nucleic Acids Res* 2020;**48**:D983–D991.

30. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* 2020;**369**:1318–1330.

31. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* 2019;**35**:4851–4853.

32. Breeze CE, Reynolds AP, van Dongen J, Dunham I, Lazar J, Neph S et al. eFORGE v2.0: updated analysis of cell type-specific signal in epigenomic data. *Bioinformatics* 2019;**35**:4767–4769.

33. Schofield EC, Carver T, Achuthan P, Freire-Pritchett P, Spivakov M, Todd JA et al. CHiCP: a web-based tool for the integrative and interactive visualization of promoter capture Hi-C datasets. *Bioinformatics* 2016;**32**:2511–2513.

34. Tse SM, Krajinovic M, Chauhan BF, Zemek R, Gravel J, Chalut D et al. Genetic determinants of acute asthma therapy response in children with moderate-to-severe asthma exacerbations. *Pediatr Pulmonol* 2019;**54**:378–385.

35. Leiter K, Franks K, Borland ML, Coleman L, Harris L, Le Souëf PN et al. Vitamin D receptor polymorphisms are associated with severity of wheezing illnesses and asthma exacerbations in children. *J Steroid Biochem Mol Biol* 2020;**201**:105692.

36. Tsai C-H, Wu AC, Chiang B-L, Yang Y-H, Hung S-P, Su M-W et al. CEACAM3 decreases asthma exacerbations and modulates respiratory syncytial virus latent infection in children. *Thorax* 2020;**75**:725–734.

37. Shrine N, Portelli MA, John C, Soler Artigas M, Bennett N, Hall R et al. Moderate-to-severe asthma in individuals of European ancestry: a genome-wide association study. *Lancet Respir Med* 2019;**7**:20–34.

38. Kong D-H, Kim Y, Kim M, Jang J, Lee S. Emerging Roles of Vascular Cell Adhesion Molecule-1 (VCAM-1) in Immunological Disorders and Cancer. *Int J Mol Sci* 2018;**19**:1057.

39. Hortelano S, López-Fontal R, Través PG, Villa N, Grashoff C, Boscá L et al. ILK mediates LPS-induced vascular adhesion receptor expression and subsequent leucocyte trans-endothelial migration. *Cardiovasc Res* 2010;**86**:283–292.

40. Lee J-H, Sohn J-H, Ryu SY, Hong C-S, Moon KD, Park J-W. A novel human anti-VCAM-1 monoclonal antibody ameliorates airway inflammation and remodelling. *J Cell Mol Med* 2013;**17**:1271–1281.

41. Kitagawa H, Shimakawa H, Sugahara K. The tumor suppressor EXT-like gene EXTL2 encodes an alpha1, 4-N-acetylhexosaminyltransferase that transfers N-acetylgalactosamine and N-acetylglucosamine to the common glycosaminoglycan-protein linkage region. The key enzyme for the chain initiation of he. *J Biol Chem* 1999;**274**:13933–13937.

42. Nadanaka S, Zhou S, Kagiyama S, Shoji N, Sugahara K, Sugihara K et al. EXTL2, a member of the EXT family of tumor suppressors, controls glycosaminoglycan biosynthesis in a xylose kinase-dependent manner. *J Biol Chem* 2013;**288**:9321–9333.

43. Pu A, Mishra MK, Dong Y, Ghorbanigazar S, Stephenson EL, Rawji KS et al. The glycosyltransferase EXTL2 promotes proteoglycan deposition and injurious neuroinflammation following demyelination. *J Neuroinflammation* 2020;**17**:220.

44. Nadanaka S, Hashiguchi T, Kitagawa H. Aberrant glycosaminoglycan biosynthesis by tumor suppressor EXTL2 deficiency promotes liver inflammation and tumorigenesis through Toll-like 4 receptor signaling. *FASEB J* 2020;**34**:8385–8401.

45. Hennessy EJ, Sheedy FJ, Santamaria D, Barbacid M, O'Neill LAJ. Toll-like receptor-4 (TLR4) down-regulates microRNA-107, increasing macrophage adhesion via cyclin-dependent kinase 6. *J Biol Chem* 2011;**286**:25531–25539.

46. Banno A, Reddy AT, Lakshmi SP, Reddy RC. PPARs: Key Regulators of Airway Inflammation and Potential Therapeutic Targets in Asthma. *Nucl Recept Res* 2018;**5**.

47. Kumari A, Dash D, Singh R. Lipopolysaccharide (LPS) exposure differently affects allergic asthma exacerbations and its amelioration by intranasal curcumin in mice. *Cytokine* 2015;**76**:334–342.

48. Hadjigol S, Netto KG, Maltby S, Tay HL, Nguyen TH, Hansbro NG et al. Lipopolysaccharide induces steroid-resistant exacerbations in a mouse model of allergic airway disease collectively through IL-13 and pulmonary macrophage activation. *Clin Exp Allergy* 2020;**50**:82–94.

49. Böhlig L, Friedrich M, Engeland K. p53 activates the PANK1/miRNA-107 gene leading to downregulation of CDK6 and p130 cell cycle proteins. *Nucleic Acids Res* 2011;**39**:440–453.

50. Yang L, Zhang B, Wang X, Liu Z, Li J, Zhang S et al. P53/PANK1/miR-107 signalling pathway spans the gap between metabolic reprogramming and insulin resistance induced by high-fat diet. *J Cell Mol Med* 2020;**24**:3611–3624.

51. El-Husseini ZW, Gosens R, Dekker F, Koppelman GH. The genetics of asthma and the promise of genomics-guided drug target discovery. *Lancet Respir Med* 2020;**8**:1045–1056.

52. Bankovic J, Stojsic J, Jovanovic D, Andjelkovic T, Milinkovic V, Ruzdijic S et al. Identification of genes associated with non-small-cell lung cancer promotion and progression. *Lung Cancer* 2010;**67**:151–159.

53. Ma J, Zhao J, Lu J, Wang P, Feng H, Zong Y et al. Cadherin-12 enhances proliferation in

colorectal cancer cells and increases progression by promoting EMT. *Tumour Biol* 2016;**37**:9077–9088.

54. Zhao J, Li P, Feng H, Wang P, Zong Y, Ma J et al. Cadherin-12 contributes to tumorigenicity in colorectal cancer by promoting migration, invasion, adhersion and angiogenesis. *J Transl Med* 2013;**11**:288.

55. Singh SR, Sutcliffe A, Kaur D, Gupta S, Desai D, Saunders R et al. CCL2 release by airway smooth muscle is increased in asthma and promotes fibrocyte migration. *Allergy* 2014;**69**:1189–1197.

Table 1. Association results for the top hit in the discovery stage and sentinel variants with significant and consistent effects in the discovery and replication phases, and the meta-analysis across both phases.

			Discovery			Replication			Meta-analysis (discovery and replication)		
ID†	rsID	Closest gene	OR (95% CI)	P	Cochran's Q P	OR (95% CI)	P	Cochran's Q P	OR (95% CI)	P	Cochran's Q P
1:101210560:C:T	rs12091010	*EXTL2*	0.82 (0.75–0.90)	9.05E-06	4,83E-01	0.89 (0.82–0.97)	5.35E-03	4.92E-01	0.86 (0.81–0.91)	4.23E-07	4.47E-01
5:22659406:T:C	rs6888198	*CDH12*	1.37 (1.23–1.54)	1.95E-08	5,82E-01	1.02 (0.90–1.15)	7.72E-01	7.18E-01	1.24 (1.05–1.45)	2.41E-06	1.89E-02
10:91376299:T:C	rs943126	*PANK1*	0.85 (0.78–0.92)	3.10E-05	8,01E-02	0.92 (0.86–0.98)	1.30E-02	3.87E-01	0.89 (0.85–0.94)	4.93E-06	7.91E-02

†The variant identifier corresponds to chromosomal position (hg19) followed by non-tested allele and tested allele. Abbreviations: 95% CI: 95% confidence interval; OR: Odds ratio; P: P-value.

Table 2. Results from the meQTL analysis in whole blood in the GALA II and SAGE studies for genome-wide significant hit in the discovery and two SNPs that were replicated.

			Mexican Americans			Puerto Ricans			African Americans			Meta-analysis				
SNP-CpG pair	Position (hg19)	Closest genes	Coef	SE	P	Coef	SE	P	Coef	SE	P	Coef	SE	P	Cochran's Q P	Storey q
rs943126-cg26800131	91574784	*KIF20B*	-0.21	0.11	**2.99E-04**	-0.07	0.04	6.99E-02	-0.07	0.04	8.87E-02	-0.18	0.10	**1.85E-06**	4.45E-03	**9.95E-04**
rs943126-cg14920044	91296311	*SLC16A12*	0.09	0.05	9.77E-02	0.24	0.06	**1.68E-05**	0.12	0.06	**4.68E-02**	0.15	0.03	**4.61E-06**	1.24E-01	**1.24E-03**
rs943126-cg20654695	91444521	*KIF20B/PANK1*	-0.10	0.08	2.12E-01	-0.06	0.04	9.04E-02	-0.15	0.03	**2.04E-05**	-0.10	0.02	**9.96E-06**	1.98E-01	**1.79E-03**
rs943126-cg25770176	91405685	*PANK1*	-0.09	0.03	**1.43E-03**	0.00	0.09	2.78E-01	-0.07	0.02	**3.00E-03**	-0.07	0.02	**1.86E-05**	1.49E-01	**2.50E-03**
rs12091010-cg05612904	101491636	*DPH5*	-0.07	0.04	6.99E-02	-0.21	0.11	**2.99E-04**	-0.09	0.04	**2.72E-02**	-0.10	0.02	**2.31E-05**	9.52E-02	**1.20E-02**
rs943126-cg00475140	91404454	*PANK1*	-0.21	0.07	**1.39E-03**	-0.13	0.08	1.05E-01	-0.19	0.09	**3.19E-02**	-0.18	0.04	**4.28E-05**	5.25E-01	**4.60E-03**
rs943126-cg15620114	91296457	*SLC16A12*	0.09	0.08	2.64E-01	0.27	0.09	**3.87E-03**	0.20	0.08	**7.76E-03**	0.18	0.05	**1.53E-04**	5.70E-01	**1.32E-02**
rs943126-cg04957662	91411382	*KIF20B/PANK1*	-0.34	0.79	6.69E-01	-1.16	0.29	**1.33E-04**	-0.77	0.26	**2.89E-03**	-1.02	0.27	**1.72E-04**	3.14E-01	**1.32E-02**

Abbreviations: Coef: Coefficient of the regression; P: P-value; SE: Standard error; Storey q: Storey q-value.

Table 3. Sensitivity analysis for rs12091010 and rs943126 in individuals from the discovery stage.

	Exacerbations in the last 6 months			Exacerbations in the last 12 months					
	European-descent populations			European-descent populations			Multi-ethnic meta-analysis		
rsID	OR (95% CI)	P	Cochran's Q P	OR (95% CI)	P	Cochran's Q P	OR (95% CI)	P	Cochran's Q P
rs12091010	0.84 (0.67-1.04)	1.08×10^{-1}	5.14×10^{-1}	0.86 (0.72-1.03)	9.13×10^{-2}	6.25×10^{-1}	0.82 (0.74-0.90)	3.45×10^{-5}	4.84×10^{-1}
rs943126	0.78 (0.64-0.96)	2.13×10^{-2}	8.61×10^{-1}	0.88 (0.74-1.04)	1.26×10^{-1}	8.28×10^{-1}	0.85 (0.78-0.93)	3.29×10^{-4}	7.50×10^{-2}

Abbreviations: 95% CI: 95% confidence interval; OR: Odds ratio; P: P-value.

FIGURES LEGENDS

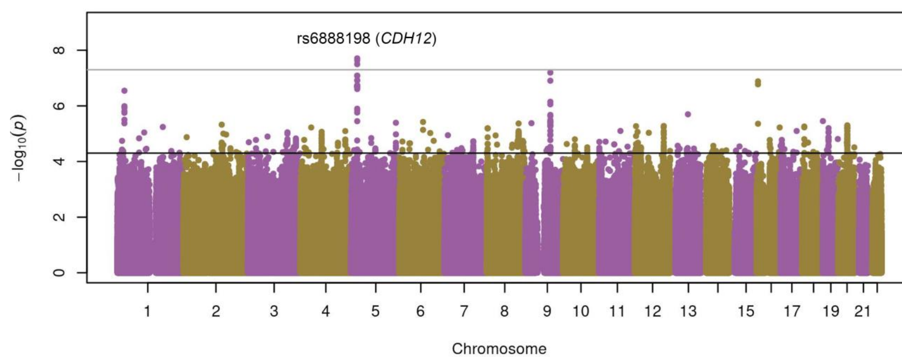


Figure 1. Manhattan plot of the results of the discovery stage of the multi-ancestry meta-analysis of GWAS of asthma exacerbations (represented as -$\log_{10}$ p-value on the y-axis) along the chromosome position of the variants analyzed (x-axis). The suggestive ($p=5x10^{-5}$) and genome-wide ($p=5x10^{-8}$) significance thresholds are indicated by the black line and dark gray lines.

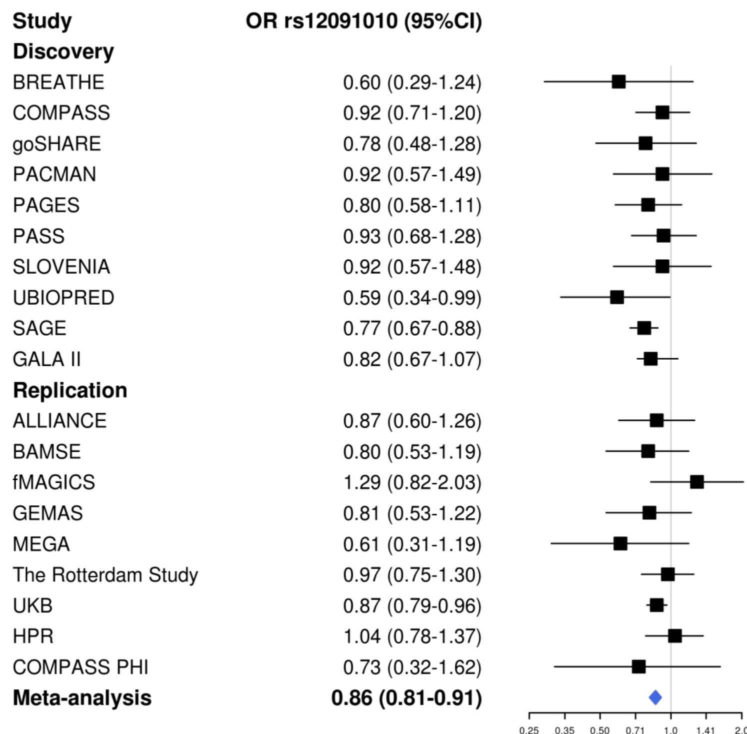


Figure 2. Forest plot of the association results for rs12091010 (*VCAM1/EXTL2*) in the meta-analysis of GWAS of asthma exacerbations.

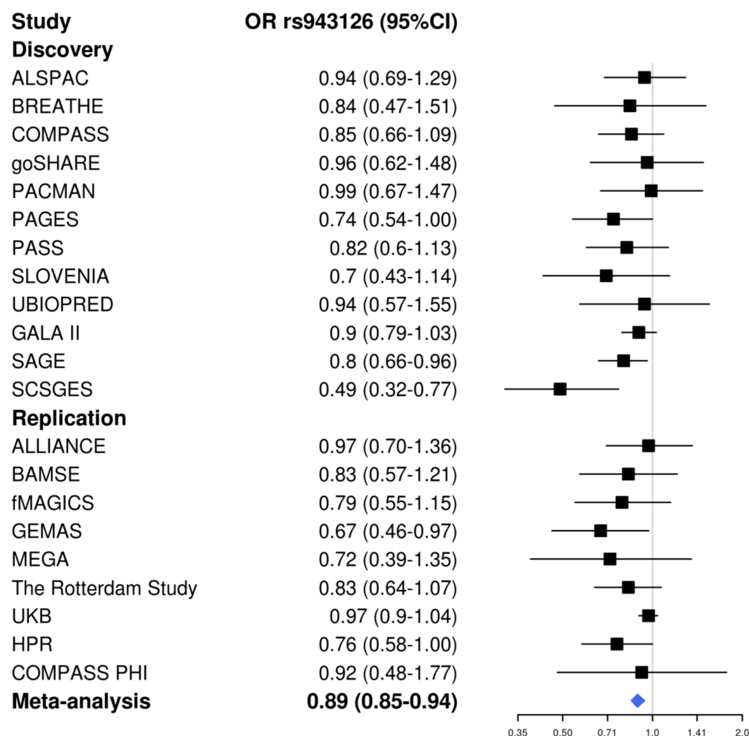


Figure 3. Forest plot of the association results for *rs943126* (*PANK1*) in the meta-analysis of GWAS of asthma exacerbations.

Link to Supporting Information:

https://drive.google.com/drive/folders/1yZry2IsJrHkHQ07EPhNO5wCEPCUAEGZ3?usp=sharing

5

General discussion

This doctoral thesis addresses the complex pathophysiology of AEs from a genetic perspective via multiple genomic approaches. To achieve this objective, we have i) reviewed all the main findings of the genetic and genomic studies of AEs published up to 15th December 2018 ii) conducted a candidate-gene study of AEs in ethnically diverse populations that assessed six genomic regions harbouring genes whose gene expression predicts future AEs iii) explored the genetic markers involved in asthma with severe exacerbations in children and youth from ethnically diverse populations using a specific phenotyping strategy; iv) leveraged local ancestry to identify chromosomal regions where genetic ancestry is associated with AEs in Hispanic/Latino children and youth with asthma and flagged the genetic loci responsible for the admixture mapping peaks; v) extended the study of genetic variation associated with AEs using a multi-ethnic framework. Taking into consideration the previous *Chapters*, here we will discuss the main findings, methodological challenges, and future directions in the genomics of AEs.

5.1 Insights into the genetics and genomics of asthma exacerbations

Chapter 1 revised all published genetic association studies of AEs until the 15th December 2018, at the beginning stage of this thesis, to identify genetic variants that could be evaluated for replication in independent studies. This search was extended up to the 15th November 2021 in *Chapter 5*. Our literature search revealed that most of these associations were detected following candidate gene approaches focused on genes previously implicated in asthma susceptibility (such as the *ADRB2* gene or sev-

eral genes located in the 17q21 asthma susceptibility locus) (Herrera-Luis et al., 2019). Some of these hypothesis-driven studies have addressed interaction with environmental exposures, such as dust mite exposure for the *IL9* (Sordillo et al., 2015), *IL10* (Hunninghake et al., 2008) and *TGFB1* (Hunninghake et al., 2008) genes or endotoxin exposure for the *LY96* gene (Kljaic-Bukvic et al., 2014). Moreover, other studies focused on the genetic variation in response to treatment in relevant asthma-related genes, such as the *FCER2* gene in ICS response (Koster et al., 2011; Tantisira et al., 2007). However, the clinical utility of these findings is hampered by their modest effect sizes, the lack of assessment for replication in independent populations of the same ethnic group or diverse ethnic backgrounds. Although two candidate-gene association studies investigated microsatellites (Hirai et al., 2018) or insertion/deletion polymorphisms (Ortega et al., 2014), genetic variation other than single nucleotide polymorphisms (SNPs) has been ignored. In addition, many of these studies exhibited modest sample size and/or low genome coverage.

The first GWAS of AEs was published in 2012 and specifically focused on the interaction of genetic variation and vitamin D levels on asthma-related hospitalizations and/or emergency room visits (Du et al., 2012). Another genome-wide scan focused on age-by-genotype interactions on AEs despite ICS use and revealed two significant associations at genome-wide level (Dahlin et al., 2020). However, it is striking that only two genome-wide scans from the nine studies published at the date of the writing, apart from the results described in this doctoral thesis, assessed genotype-environment interactions on AEs, despite the recognized crucial interplay of genetic and environmental factors on AEs (Martin et al., 2020; Ramsahai et al., 2019; di Palmo et al., 2021). On the same line, most of the pharmacogenomic studies of AEs have detected genetic associations that did not reach genome-wide significance either in asthma patients under ICS (Dahlin et al., 2020; Hernandez-Pacheco et al., 2019, 2021) or LABA treatment (Slob et al., 2021). Interestingly, one of these studies identified trichostatin A as a potential novel therapeutic target on AEs by gene-set enrichment analysis on drugs, and these findings were supported by the evidence of role of trichostatin A in asthma-related processes (Hernandez-Pacheco et al., 2021). However, this analytical approach has not been applied extensively in GWAS of AEs.

One of the ongoing challenges for genomic research is increasing the ethnic diversity (Mills and Rahal, 2020). The results from our literature search for genetic loci for AEs emphasized the lack of studies focusing on the ethnic-specific genomic predictors of AEs in African-descent or Asian populations alone. Further efforts should be made to bring together clinical and genetic data in order to reduce these ethnic disparities in genomic research. However, the identification of ethnic-specific genomic predictors

of AEs in these populations remains a challenge due to the modest sample size of the available studies.

Another key point in genomic research is unraveling the regulatory effects of the revealed novel associations. Indeed, the investigation of functional effects of genetic variants identified is absent of many of the studies conducted to date and should be incorporated routinely in the experimental designs either by using *in vivo*, *ex vivo* and *in vitro* analyses, and/or *in silico* approaches using publicly available online resources, e.g., Kundaje et al. (2015); Schofield et al. (2016); Staley et al. (2016); Ward and Kellis (2016); Dayem Ullah et al. (2018); Kamat et al. (2019); Moore et al. (2020); GTEx Consortium (2020); Zheng et al. (2020); Ghoussaini et al. (2021); Min et al. (2021). Moreover, due to the intricacies in the interplay of risk factors, triggers, and predisposing factors (Ramsahai et al., 2019; Martin et al., 2020; di Palmo et al., 2021), sensitivity analysis should be conducted to confirm the robustness of the associations to potential confounders and assess their role in modulating the genetic association via interaction analyses.

In addition, we observed that a reduced number of candidate-gene studies and none of the GWAS of AEs published to date have assessed rare genetic variation via NGS techniques. Although this type of studies could be relevant in all ethnic groups, they are particularly interesting in minority populations that show differential rates of AEs. Needless to say, these approaches should be in hand with larger sample sizes in order to identify relevant associations in populations that are inherently complex from a genetic perspective (Schoettler et al., 2019). Despite the reduction of costs in the implementation of WGS in research, imputation-based GWAS is still the most cost-efficient genome-wide approach to assess both common and low-frequency variation (Quick et al., 2020). In this area, the recent availability of the The Trans-Omics for Precision Medicine (TOPMed) imputation reference panel (Taliun et al., 2021), including 53,831 genomes from ethnically diverse populations worldwide, is a promising strategy to increase to provide insights into the genetic architecture of human traits and diseases.

5.2 Prioritization of *DNASE1L3* in asthma exacerbations

In *Chapter 2*, we described the results of a candidate-gene study of severe AEs in minority populations with asthma for six genes whose sputum expression differentiated airway asthmatic inflammatory phenotypes (Baines et al., 2014; Fricker et al., 2019; Frøssing et al., 2021), corticosteroid treatment responsiveness (Berthon et al.,

2017), and showed high accuracy for the prediction of AEs (Fricker et al., 2019). In the discovery stage, we evaluated six genomic regions including the *ALPL*, *CLC*, *CPA3*, *CXCR2*, *DNASE1L3*, and *IL1B* genes. One polymorphism in the *DNASE1L3* region exceeded the Bonferroni-like corrected threshold of significance for the association with severe AEs in African Americans. Specifically, the C allele of rs67622929 was significantly associated with higher odds of having severe AEs. Moreover, none of the genetic variants in the other five genomic regions showed significant associations with the exacerbation status. In the replication phase, the SNP rs67622929 showed significant, and consistent effects in Hispanics/Latinos and the risk allele was associated with 1.3 times increased risk of having severe exacerbations in the meta-analysis across both populations. Notwithstanding the fact that the association of rs67622929 with severe AEs was significant and consistent in African Americans and Hispanics/Latinos, the association was not significant ($p>0.05$) in Singaporean Chinese and Europeans, as described in *Chapter 5*.

The *in silico* functional evidence suggests that rs67622929 has regulatory effects over whole-blood DNA methylation of several CpG sites located nearby *DNASE1L3* (Staley et al., 2016; Kamat et al., 2019; Min et al., 2021) and modulates *DNASE1L3* expression in several tissues, including asthma-relevant tissues, by means of eQTLs, and/or splicing quantitative trait loci. In addition, rs67622929 is associated with acetylation of the lysine residue at N-terminal position 27 of the histone H3 in the lung according to Roadmap data (Kundaje et al., 2015) within the SNPnexus annotation framework (Oscanoa et al., 2020). Moreover, *DNASE1L3* is associated with promoter/enhancers that regulate multiple transcription factor binding sites that regulate gene expression in several tissues and cells, including lung, blood cells, and macrophages, according to GeneHancer (Fishilevich et al., 2017). Although DNASE1L3 is mainly produced by macrophages (Sisirak et al., 2016), it is also found circulating in blood (Koyama et al., 2016).

DNASE1L3, located at 3p14, encodes for a member of the deoxyribonuclease I family that hydrolyses both uncomplexed and complexed DNA. DNASE1L3 mediates the breakdown of the DNA both in apoptosis (Shi et al., 2017) and necrosis (Koyama et al., 2016), and the extracellular degradation of the neutrophil extracellular traps (NETs) (Engavale et al., 2021). NETs constitute a defensive mechanism released by neutrophils that allows entrapping and killing pathogens. However, the regulation of the homeostasis of NETs is essential since the aggregation of NETs or their sustained presence in circulation could alter the inflammatory and immune activities (Engavale et al., 2021). Likewise, DNASE1L3 participates in cytokine secretion, including IL-1β release by activating the NLRP3 and NLRC4 inflammasome (Shi et al., 2017). In turn, IL-1β

promotes the production of mediators involved in the airway remodelling, such as metalloproteinases *MMP9* and *MMP12* (Lappalainen et al., 2005). Furthermore, the involvement of *DNASE1L3* in autoimmunity is highlighted by the association of genetic variation with several rheumatic diseases (Eyre et al., 2012; Mayes et al., 2014; Terao et al., 2017; Márquez et al., 2018; Acosta-Herrera et al., 2019) and the reduced serum DNASE1L3 levels in dermatomyositis/polymyositis and systemic lupus erythematosus (Zhao et al., 2017). DNASE1L3 reduced levels could cause a deficit on clearance in NETs after pathogen elimination that promotes chronic inflammation (Engavale et al., 2021). Despite this, we found that the risk allele for AE was associated with a higher *DNASE1L3* gene expression in several tissues. This is in concordance with previous studies about the effect of *DNASE1L3* gene expression in the prediction of AEs (Fricker et al., 2019). All these findings emphasize the complexity of the *DNASE1L3*-mediated homeostasis processes and the need for further research on the role of *DNASE1L3* in asthma.

Moreover, rs67622929 interacted with *FLNB* in naive CD8+ T cells through promoter capture Hi-C, according to the Open Targets Genetics database (Ghoussaini et al., 2021). *FLNB* encodes for Filamin B, which mediates the intracellular communication and signal transmission to the cytoskeleton in order to regulate the adhesion to the extracellular matrix (Xu et al., 2017), which could explain the association of several skeletal disorders with pathogenic mutations in *FLNB* (Xu et al., 2017). *FLNB* is known to participate in muscle cell differentiation, morphogenesis, chondrocyte homeostasis, and angiogenesis (Xu et al., 2017). In human umbilical endothelial cells, the small interfering RNA knockdown of *FLNB* decreased cell migration and reduced angiogenesis promoted by the vascular endothelial growth factor (VEGF) (Del Valle-Pérez et al., 2010). Interestingly, VEGF self-regulates its trafficking capability via the monoubiquitination of filamin B (Su et al., 2013).

5.3 *LINC01913* and *PKDCC* as potential targets in asthma exacerbations

The first GWAS of asthma with severe exacerbations in ethnically diverse children and youth, presented in *Chapter 3*, describes for the first time the association of a genetic variant located 87.8 kb away from the 5'-untranslated region of the *LINC01913* gene and asthma with severe exacerbations in Hispanics/Latinos and African Americans. In the meta-analysis across both ethnic groups, the SNP rs4952375 was associated with a 39% increase in the odds of having asthma with exacerbations compared

to subjects without asthma and a 25% increase in the odds of having AEs among asthma patients.

The analysis of ethnically diverse children was motivated by the lack of representation of minority populations in previous genetic studies of AEs (Herrera-Luis et al., 2019). Interestingly, this bias of ethnic diversity could be considered almost endemic in genomic research. At the time of writing, the diversity in the total GWAS participants is lowest in American admixed and African-descent populations: European (84.2%), Asian (11.5%), Other/Mixed (2.1%), Hispanic or Latin American (0.8%), African American or Afro-Caribbean (0.7%) and African (0.3%) (GWAS Diversity Monitor, 2020; Mills and Rahal, 2020). In this regard, in recent years, great efforts have been made to bridge the gap on the knowledge about genetic variation involved in AEs in ethnically diverse populations. This doctoral thesis is part of those efforts along with other studies in Hispanics/Latinos and/or African Americans (Hernandez-Pacheco et al., 2019; Yan et al., 2021a). This lack of ethnic representation in genomic research is partly explained by the reduced number and smaller sample sizes of studies including these populations compared to those focused on Europeans. Those limitations difficult the discovery of genetic loci in populations that are inherently genetically heterogeneous (Schoettler et al., 2019). Partially because of this, we opted for a specific phenotyping approach where asthma patients with severe exacerbations were compared to non-asthma subjects. Specific or extreme phenotyping approaches have been successfully applied before in the search of genetic loci involved in several disorders or traits with a genetic basis (Amanat et al., 2020), including asthma with recurrent exacerbations in young Danish children (Bønnelykke et al., 2014). In our analysis, this strategy implicated an increase of 34.1% in sample size in the discovery stage compared to an analysis restricted to asthma patients with exacerbation data. For a prevalence of AEs of 30.5%, based on sample size alone, we had a 13.8% increase in statistical power to detect associations with MAF of 23% and relative risk of 1.30 in the discovery stage compared to the analysis including only asthma patients. The downside of this approach is that some of the significant associations may be due to the asthma vs. non-asthma design rather than associations for AEs *per se*. This fact motivated us to examine the association of the genetic variants with severe AEs among asthma patients. While the 17q12-q21 locus showed non-significant association in the case-only meta-analysis comparing asthma patients with and without exacerbations, the SNP rs4952375 remained significant, suggesting its genuine involvement in AEs.

The standard threshold of genome-wide significance ($p<5x10^{-8}$) could considered stringent in admixed American populations due to the variation in LD patterns across ancestral groups (Kanai et al., 2016). Based on the 1KGP data, the genome-wide sig-

nificance in admixed American populations was established as $p<1.83\times10^{-7}$ (Kanai et al., 2016). Although the less stringent threshold was applied in the meta-analysis, no additional loci with significant and consistent effects would have been revealed as genome-wide significant.

We attempted to validate the association of rs4952375 and AEs in additional populations. However, this SNP was not available in Singaporean Chinese probably due to low frequency in East Asians according to the 1KGP (MAF= 0.009%) and showed non-significant association with AEs in European patients with asthma ($p>0.05$), as described in *Chapter 5*.

The assessment of functional effects flagged rs4952375 as an eQTL for the expression of *LINC01913* in lung tissues. *LINC01913*, located at 2p21, encodes for a long intergenic non-protein coding RNA, which is predominantly expressed in lung, according to the Genotype-Tissue Expression (GTEx Consortium, 2020). Although the biological function of *LINC01913* is unknown, it likely participates in the regulation of relevant processes, as described for other long intergenic non-protein coding RNAs in asthma (Wang et al., 2019). On the other hand, *PKDCC*, also known as *VLK* or *SGK493*, encodes for a tyrosine-protein kinase secreted to the extracellular matrix that participates in signal transduction (Bordoli et al., 2014). Specifically, PKDCC mediates the phosphorylation of the endoplasmic reticulum chaperone ERP29 and multiple matrix metaloproteinases, such as MMP1 (Bordoli et al., 2014). MMP1 promotes airway smooth muscle proliferation and extracellular matrix remodeling, and its activation state has been associated with the severity of AEs (Naveed et al., 2017).

The SNP rs4952375 was also a meQTL for a CpG site nearby *PKDCC*. *PKDCC* has diverse biological roles, including the phosphorylation of extracellular matrix proteins that mediate cell adhesion in trabecular meshwork cells (Maddala et al., 2017) and axon guidance in retinal ganglion cells (Harada et al., 2019), regulation of the endoplasmic reticulum proteostasis (Farhan, 2020) and Hedgehog signaling (Kim et al., 2020), and platelet aggregation and granule release (Revollo et al., 2021). Interestingly, genetic variants from *PKDCC* have been associated with skeletal development and homeostasis, lung function and atopy, according to the GWAS catalog (Buniello et al., 2019). However, in our study, rs4952375 and previous loci *PKDCC* described in the GWAS catalog (Buniello et al., 2019) were not associated with predicted lung function or atopy. In addition, evidence from mice indicates *PKDCC* involvement in embryogenesis, lung development, morphology, and cyanosis, as described in the Open Targets Genetics database (Ghoussaini et al., 2021). Still, further research is needed to confirm and disentangle the specific biological function of *LINC01913* and *PKDCC* in the pathogenesis of AEs.

5.4 Leveraging local ancestry to detect genetic signals

Leveraging local ancestry to identify genetic variants common in an admixed populations that show differential rates of a specific trait or distinct patterns of relevant quantitative phenotypes has successfully uncovered ethnic-specific and population-shared loci. In the context of asthma, this approach has been applied in Hispanics/Latinos, African Americans, and southwestern Europeans (Galanter et al., 2014; Daya et al., 2019; Gignoux et al., 2019; Guillen-Guio et al., 2020) and have also uncovered loci for relevant traits such as IgE levels in Latinos (Pino-Yanes et al., 2015a), lung function in Puerto Ricans, African Americans and African-descent populations (Parker et al., 2014; Fonseca et al., 2020; Lee et al., 2020b), or even smoking-dependent loci of lung function in African Americans via local ancestry–smoking interactions analysis (Ziyatdinov et al., 2020). African American populations are the results of two-way admixture events between European and African ancestries. In contrast, the genomes of Hispanic-Latinos are mosaics of chromosomal regions resulting from the admixture among Europeans, indigenous populations from America and African ancestries. In Hispanic/Latinos, admixture proportions vary significantly by geographical region based on the specific historic events that originated the admixed population (Bryc et al., 2015). For example, the African component is significantly higher in Puerto Ricans than in Mexican-descent populations, where the Indigenous American proportion is much higher (Galanter et al., 2014). Notably, asthma prevalence is higher in Puerto Ricans and Mexicans, which could be partially explained by the direction of the effects of genetic ancestry on asthma susceptibility. Among these ethnic subgroups, African ancestry is associated with higher odds of asthma, whereas Indigenous American ancestry is associated with a lower risk of asthma (Pino-Yanes et al., 2015b).

In *Chapter 4*, we reported the results of the first admixture mapping of AEs in Hispanics/Latinos and revealed two novel genetic signals for AEs at chromosome 5q32-q33.1. The importance of this genomic region is highlighted by the fact that it has been involved in asthma by linkage studies in families, and it is located near several relevant genes previously implicated in asthma susceptibility and asthma treatment response (e.g., *IL4*, *IL3*, *IL5*, *IL9*, *L13*, *RAD50*, *PTGER4*, *TSLP*, or *ADBR2*) (Ober and Hoffjan, 2006). Despite the fact that Latinos exhibit differential rates of AEs compared to other ethnic groups (Akinbami et al., 2012; Oraka et al., 2013; Akinbami et al., 2014; Rosser et al., 2014), we did not find a significant association between global genetic ancestry and severe AEs defined based on acute asthma care, hospitalizations, or use of OCS in the previous year, in Puerto Ricans and Mexican Americans, in line with previous evidence (Brehm et al., 2012). In this regard, the following question arises: Does

the fact that admixed populations exhibit differential rates of a specific trait but no significant association between the trait of interest and global genetic ancestry hampers the application of admixture mapping? Although no significant association with global genetic ancestry is detected, if the population shows a high prevalence of the trait, it is still possible that certain regions of the genome are relevant for the trait. As discussed before, both genetics and environmental factors interplay to increase the odds of AEs and evidence of the role of gene-environment interactions in AEs have been previously published (Du et al., 2012; Bønnelykke et al., 2014). Therefore, it is likely that environmental factors interact with specific regions to modulate the association of local ancestry with AEs, as has been previously shown for smoking and lung function (Ziyatdinov et al., 2020).

Our analysis revealed distinct genomic regions associated with AEs in Puerto Ricans and Mexican Americans. Fine-mapping of these regions flagged two significant associations in the 5q32 locus in Mexican Americans that showed significant and consistent effects in Hispanic/Latino sub-groups. When these two genetic variants were assessed for replication in non-Hispanic/Latino populations, no significant association with AEs was found in Singaporean Chinese, African Americans, and Europeans. In an effort to disentangle if a set of environmental and socioeconomic factors may module the effect of the association, we confirmed the robustness of the association of AEs and the identified variant by accounting for second-hand smoking exposure, insurance status, income quartile, maternal education, and obesity. Our results indicated that this compendium of factors did not play a role in modulating of the genetic association, since the effects of the association for the genetic variants with AEs remained consistent after adjustment by potential confounders.

The meQTL analysis revealed these two genetic loci associated with whole-blood DNA methylation levels at several CpG sites at the *DYPSL3* region. *DYPSL3*, also known as *CRMP4*, participates in the regulation of epithelial-to-mesenchymal transition in several cancer types (Matsunuma et al., 2018; Tsai et al., 2020b). Specifically, *DYPSL3* reduces the expression of Snail and Twist, key transcription factors in epithelial-to-mesenchymal transition in breast cancer cells. However, these transcription factors can also induce *DYPSL3* expression (Matsunuma et al., 2018). However, *DYPSL3* may play different roles in cell migration and regulation of metastasis per cancer types, according to the previous studies (Tsai et al., 2020b). In neuronal tissues, *DYPSL3* is a signal transductor involved in axonal regeneration (Girouard et al., 2020).

In addition, we observed that the rs1144986 G allele associated with lower odds of having AEs increased *SCGB3A2* gene expression in lung tissues. The scarce evidence of the role of *SCGB3A2* provides limited information about its involvement in

non-canonical inflammation, lung clearance of pathogens, and anti-fibrotic processes. In lung cancer cells, it joins to syndecan-1 to prompt the intake of both SCGB3A2 and LPS. As a response to the cytosolic presence of LPS, the non-canonical caspase-11 inflammasome is activated, and subsequently, cancer cell death via pyroptosis is induced (Yokoyama et al., 2018, 2021). Similar to our results, a higher gene expression of *SCGB3A2* was linked to a protective effect. Specifically, Yokoyama et al. (2021) found that patients with cancer with higher expression exhibited higher survival rates. This protective mechanism in cancer may be a residual ramification of the caspase-11-induced host defense mechanism against pathogens. Caspase-11 provides protection against Gram-negative bacterial pathogens. This activity is in a delicate balance with the promotion of airway inflammation induced by caspase-11 hyper-activation (Oh et al., 2020). In bronchial epithelial cells, the downregulatory effect of SCGB3A2 over LPS-triggered airway inflammation is accompanied by a reduction in *TNF-α* and *CXCL8* gene expression (Wang et al., 2015). Additionally, in the same cell type, SCGB3A2 competes with LPS to join the macrophage scavenger receptor with collagenous structure (MARCO), which participates in the defense against pathogens in the airways (Bin et al., 2003).

The anti-fibrotic effect of *SCGB3A2* in the lung is thought to be exerted via downregulation of the TGFβ1 pathway. TGFβ1 is a profibrotic cytokine secreted by airway fibroblasts, myofibroblasts, inflammatory cells, and bronchial epithelial cells. TGFβ1 binds to its receptor in the extracellular membrane, which triggers the cytosolic phosphorylation of SMAD2/3. This phosphorylated complex joins to SMAD4 to translocate to the nucleus, where it triggers fibroblast to myofibroblast transition, activates the production of epithelial cell-secreted inhibitors of metalloproteinase, and inhibits the secretion of matrix metalloproteinases (Royce et al., 2012). Specifically, SCGB3A2 promoted IFNγ-independent phosphorylation of STAT1 and increased SMAD7 expression, and inhibits SMAD2/3 phosphorylation, which in turn inhibits phosphorylation of SMAD2/3 (Kurotani et al., 2011). All this together along with the fact that IL-4, IL-5, and IL-13 levels are reduced in ovalbumin-induced mice pre-treated with *SCGB3A2* (Mandal et al., 2004) and IL-5 reduces *SCGB3A2* gene expression (Chiba et al., 2005) suggest that *SCGB3A2* merits further consideration as a potential therapeutic target in asthma. Moreover, our results adds to the growing evidence of the role of secretoglobins in diseases of the airways (Mootz et al., 2021).

Despite previous evidence of the association of severe AEs and African ancestry in African Americans (Rumpel et al., 2012; Grossman et al., 2019), in African Americans from the SAGE study, no significant association between severe AEs and global African ancestry was found for global ancestry estimated either from genome-wide genotyp-

ing data or WGS data, as described in *Chapter 4*, respectively. In addition, results remained non-significant in sex-stratified analysis and an analysis focused on individuals with African ancestry equal or over the median value (data not shown). In terms of the local ancestry, we identified significant admixture mapping peaks for African ancestry at chromosome 3q26.31, 3q26.33, 3q27.3-q28, 3q28-q29 (data not shown). For this analysis, we followed the same procedure described in *Chapter 4* except that local ancestry was calculated using a two-way ancestral model based on the CEU and YRI references. The subsequent fine-mapping of common and rare SNPs within the significant admixture mapping peaks did not reveal any significant association that accounted for the admixture mapping peaks (data not shown). In this regard, it is worth noting that the effect of the association of local African ancestry with AEs was in the opposite direction to the one previously described for global ancestry in African Americans (Rumpel et al., 2012; Grossman et al., 2019). Recently, a prospective study among African American children with asthma in the United States revealed that the association of African ancestry and asthma re-admissions was mediated by disease management, socio-economic factors, and the interaction of those two factors and indoor exposures (Mersha et al., 2021). This is particularly interesting due to the evidence suggesting that the reduction in in-person interactions may have modulated the decrease of AEs in African Americans after the lockdown occasioned by the COVID-19 pandemic (Salciccioli et al., 2021). All this together suggests that further genetic epidemiological studies that comprise extensive assessment of socio-economic and environmental factors are required to confirm these associations and untangle the role of the genetic ancestry-environment interactions in AEs.

5.5 *VCAM1/EXTL2* and *PANK1*: novel loci for asthma exacerbation

Excluding the work described in this doctoral thesis, a total of four studies have assessed genome-wide genetic variation involved in AEs independently of treatment or gene-environment interactions. Specifically, Bønnelykke et al. (2014) investigated early childhood asthma with recurrent hospitalizations, McGeachie et al. (2015) assessed children with asthma, Yan et al. (2021a) focused on Hispanics/Latinos and Yan et al. (2021b) analyzed White British adults. However, none of the studies have leveraged clinical and genome-wide genetic variation in a multi-ancestral setting to identify genetics signals for AEs.

The first multi-ancestral meta-analysis of GWAS of AEs, described in *Chapter 5*, assessed 9.6 million genetic variants in individuals with asthma from populations worldwide, including predominantly European-descent individuals, Hispanics/Latinos, Singaporean Chinese, and African Americans. Despite the fact that the *CDH12* locus surpassed the genome-wide significance in the discovery stage, this association was not significant ($p<0.05$) in the replication stage. Among the independent polymorphisms suggestively associated with AEs, two were replicated in the second stage. These two genetic variants were located in the intergenic region of the *VCAM1/EXTL2* gene and the intronic region of *PANK1*.

VCAM1 encodes a sialoglycoprotein predominantly expressed in the endothelial cells' luminal and lateral surface and is implicated in several cancer types and immunological disorders (Kong et al., 2018). *VCAM1* expression is regulated by multiple factors, such as cytoskeletal alignment, shear stress, and pro-inflammatory cytokines (Lee and Rosenberg, 2013; Kong et al., 2018; Fallon and Hinds, 2021). Specifically, IL-4, IL-13, TNF-α, eotaxin, or LPS have been shown to increase *VCAM-1* expression during eosinophil recruitment (Hortelano et al., 2010; Lee and Rosenberg, 2013; Kong et al., 2018). After VCAM1 binds its eosinophil receptor a4β1 integrin binding, it promotes lung eosinophilia by activating a molecular cascade including Ca^{2+} and RAC1 (Kong et al., 2018). RAC1 promotes two different processes: the formation of actin stress fibers (Wittchen, 2009) and disarranging of intercellular junctions via the activation of nicotinamide adenine dinucleotide phosphate oxidase 2 (NOX2) (Kong et al., 2018). In turn, NOX2 promotes the production of endothelial cell matrix metalloproteinases that enhance the vascular passageway opening for transendothelial migration, and increases the production of protein kinase Ca, and protein tyrosine phosphatase 1B (Deem and Cook-Mills, 2004; Lee and Rosenberg, 2013; Kong et al., 2018). Moreover, fine ambient particle matter has been shown to increase *VCAM1* expression and promote polymorphonuclear leukocyte infiltration through the endothelial-alveolar epithelial barrier, which leads to airway inflammation and injury (Cui et al., 2019). Direct blocking of VCAM1 via antibodies impairs eosinophils recruitment, eosinophilic inflammation, and airway hyperresponsivenes (Kong et al., 2018). In the same line, blocking the NF-κB signalling inhibits VCAM1 secretion and decreases the effects of LPS-induced airway epithelial cell injury (Sun et al., 2019)

EXTL2 encodes a protein that participates in the quality control of glycosaminoglycans (GAGs) biosyntesis. Specifically, the protein encoded by *EXTL2* transfers N-acetylgalactosamine and N-acetylglucosamine to the tetrasaccharide linkage region of the GAG chain, which is suggested to stop chain elongation (Kitagawa et al., 1999). Thus, decreased/inhibited *EXTL2* leads to over-accumulation of GAGs (Nadanaka et al.,

2013) with Toll-like receptor 4 agonist capacity (Nadanaka et al., 2020). This GAGs over-accumulation can also enhance the production of inflammatory and extracelular matrix remodeling mediators (Pu et al., 2020). Katta et al. (2015) showed that a small interfering RNA (siRNA) knockdown of *EXTL2* in human embryonic kidney cells increased GAG chain length. On the other hand, a siRNA in fibroblasts reduced heparan sulfate proteoglycans synthesis in disorders associated with mucopolysaccharidoses production (Kaidonis et al., 2010; Canals et al., 2015). The fact that GAGs constitute targets for the attachment of pathogens to host cells (Shi et al., 2021), along with the conflicting evidence about the role of *EXTL2* on GAGs synthesis, highlights the need for further studies about *EXTL2*.

The protein encoded by the *PANK1* gene is the first enzyme participating in the coenzyme A (CoA) biosynthesis. PANK1 is regulated by the transcription factor peroxisome proliferator-activating receptor α (PPAR-α) (Ramaswamy et al., 2004; Hennessy et al., 2011). PPAR-α exhibits anti-inflammatory activity by reducing the production of pro-inflammatory mediators (i.e., TNF-α) and activation of anti-inflammatory regulators (e.g., IL-10) (Banno et al., 2018). PPAR-α deficiency increases airway hyperresponsiveness and eosinophilia, and reduced the anti-inflammatory effect of dexamethasone in lung from mice (Banno et al., 2018). Moreover, PPAR-α controls the expression of miR-107, encoded within the intron 5 of *PANK1*, which participates in the regulation of macrophage adhesion in response to LPS (Hennessy et al., 2011). Specifically, gene expression of both *PANK1* and miR-107 was shown to be reduced in response to LPS (Hennessy et al., 2011).

All this evidence suggests that these genes may be implicated in key asthma processes. Our analysis of functional effects revealed that rs12091010 (*VCAM1*/*EXTL2*) and rs943126 (*PANK1*) exerted regulatory effects over DNA methylation in several CpG sites located nearby both polymorphisms according to the meQTL results in Latinos/Hispanics and African Americans. We also aimed to prioritize gene targets via eQTL analysis, where rs12091010 and rs943126 were shown to regulate gene expression of *EXTL2* and *PANK1*, respectively, in publicly available data from Europeans (Kamat et al., 2019). However, these results require additional validation since the association was not replicated in the GTEx data (GTEx Consortium, 2020).

In terms of the ethnic-stratified analysis, we investigated if the polymorphisms associated at a suggestive level in the meta-analysis of GWAS in predominantly European descents showed significant and consistent effects in Europeans from the replication stage. However, none of the SNPs was replicated consistently across studies (data not shown). The stratified results from the other ethnic groups were not evaluated due

to the fact that exclusively one study per non-European ancestral group analysis was available in the discovery.

We also assessed for associations with AEs 47 previous associations for AEs (Herrera-Luis et al., 2019; Tse et al., 2019; Leiter et al., 2020; Tsai et al., 2020a; Yan et al., 2021a,b) or moderate-to-severe asthma (Shrine et al., 2019). Although we found evidence of association for several loci at p<0.05 in Europeans (*GSDMB*, *RAD50*, *HLA-DQB1*, *IKZF3*, and *WDR36*), Latinos (*HLA-DQB1*, *ADAM33*, and *SMAD3*), African Americans (*CDHR3*, *VDR*, *MUC5AC*, *SMAD3* and *C11orf30*) and Singaporean Chinese (*TSLP*), none of these surpassed a stringent threshold for significance corrected for multiple testing. Two of these loci were consistently associated at p<0.05 in more than one ethnic group (*SMAD3* and *HLA-DQB1*). Interestingly, both polymorphisms exhibit pleiotropic effects. For example, rs72743461 (*SMAD3*) is associated at genome-wide level with several traits, such as eosinophil counts (Sakaue et al., 2021), allergic disease (Johansson et al., 2019), or medication use (Sakaue et al., 2021). Likewise, the SNP rs9275503 at *HLA-DQB1* is associated at genome-wide level with cell blood counts (Astle et al., 2016), immunoglobulin A deficiency (Bronson et al., 2016), or rheumatoid arthritis (Okada et al., 2014), among other traits.

5.6 Study limitations

Although this doctoral thesis has provided some insight into the genetic variation of AEs, we acknowledge that this thesis also has a number of limitations. First, although we defined AEs based on three different variables (asthma acute care, asthma-related hospitalizations, and/or OCS use) following the ERS/ATS criteria (Reddel et al., 2009), our definition was based on retrospective self-reported data in most of the studies, which could be biased due to the patient's subjectivity. Likewise, we were unable to assess the association of genetic variants with the annual number of exacerbations, the temporal distance among events, or the time-to-first exacerbation event due to the non-availability of data in most studies. On the same line, in *Chapter 5*, in order to reach enough sample size to increase our power to detect associations, we included studies that evaluated self-reported retrospective data about AEs at 6, 12, or 48 months, which could introduce bias due to the time-dependent probability of occurrence of exacerbations. Moreover, we analyzed data from children and adults together despite some studies suggest the existence of age-dependent phenotypes of AEs (Bønnelykke et al., 2014). However, our findings were consistent across the vast majority of the studies.

Second, asthma severity was defined using a proxy based on pharmacological stepwise approaches following the guidelines that were on use at the time of the recruitment for each study and not the current GINA guidelines (Global Initiative for Asthma, 2021a) given that current guidelines would not easily extrapolate to the context of period recruitment.

Third, none of the genetic loci that showed significant and consistent effects across the discovery and replication stages of the multi-ethnic meta-analysis of GWAS of AEs, described in *Chapter 5* exceeded the genome-wide significance threshold of $p<5x10^{-8}$. In terms of the threshold for suggestive associations, different arbitrary cut offs have been used to identify associations signals to evaluate for replication in independent populations (Hernandez-Pacheco et al., 2019). In *Chapter 5*. we established a sub-optimal threshold for suggestive associations based on the hypothesis that this could lead to the discovery of novel loci that could be flagged in future studies of larger samples size (Hammond et al., 2021).

Fourth, although we performed sensitivity analyses to confirm the robustness of the association between the genetic signals and AEs, the lack of availability of homogeneous information about relevant environmental exposures (e.g., viral infections) in the studies included in this doctoral thesis limited the scope of the sensitivity analysis, particularly in *Chapter 5*, where the highest number of studies was included. Although the robustness of the association was confirmed by the two-stage design for the loci identified in *Chapters 2-5*, we were not able to analyze the role of relevant exposures in the modulation of the association with AEs.

Fifth, despite the fact that we identified several significant admixture mapping peaks in Puerto Ricans and Mexican Americans, we failed to identify common or rare genetic variants that accounted for the admixture mapping peaks in Puerto Ricans, despite the larger sample size. As discussed in *Chapter 4*, this may be due to the fact the significant admixture mapping peaks were associated with Indigenous American ancestry, which represents the minor contribution to the genetic background of this Latino subgroup. Still, future studies of larger sample sizes may lead to the identification of the genetic signals driving the associations of AEs and local ancestry at these genomic regions.

Sixth, although we identified significant genetic associations for AEs, exploratory analyses not included in the chapters of this doctoral thesis suggested that none of the SNPs was considered a good discriminatory marker to predict AEs based on the area under the curve (AUC) of the receiver operating characteristics (ROC) (data not shown). Although the classification model including rs28836559, rs1144986, and

clinical data showed slight improvement over the model based on clinical data (AUC [95% CI]: 0.70 [0.66-0.74] vs 0.65 [0.61-0.69], $p=1.06 \times 10^{-3}$), this observation was not replicated in Puerto Ricans (AUC [95% CI]: 0.72 [0.69-0.75] vs 0.71 [0.68-0.75], $p=1.53 \times 10^{-1}$). Moreover, although we evaluated several polygenic risk scores based on the meta-analyses described in *Chapter 5*, none of the PRS was replicated in independent populations (data not shown). This could be due to underlying heterogeneity in the GWAS summary statistics ('base dataset'), as addressed above, and/or the limited number of samples included in the 'target datasets' that could be considered for PRS estimation (i.e., studies with genome-wide genotyping data available that were not included in the meta-analyses of the discovery stage): GEMAS (n=282) and MEGA (n=138).

Seventh, although evidence of functional effects was provided using i) DNA methylation data and gene expression data from whole-blood from ethnically diverse populations, and ii) *in silico* assessment via publicly available data, and discussed the possible role of the identified genes according to the literature-based evidence, more research is needed to confirm whether the genetic variants identified are therapeutic targets of interest in the pathogenesis of AEs.

5.7 Future directions in the genomics of asthma exacerbations

Despite the advances of this doctoral thesis, further studies on the exploration of the specific role in the pathogenesis of AEs of the genes revealed by our work are needed. This would require the design of studies that include the recollection of prospective/retrospective reporting of AEs, the annual number of exacerbations, the temporal distance among events and/or the time-to-first exacerbation. Moreover, these studies should be conducted in individuals with diverse ethnic backgrounds, considering environmental exposures, and stratifying by age group. These improvements could contribute to disentangle the complex interplay between genetics, behavioural, and environmental factors. The analysis of specific sub-types and homogeneous groups of asthma patients could decrease clinical phenotype heterogeneity and uncover novel relevant loci without requiring larger sample sizes.

Our literature mining of genetic loci for AEs revealed that most of the studies conducted to date followed candidate-gene strategies and focused on SNPs. Moving forward precision medicine, conducting candidate-gene studies should be carefully considered in light of their weaknesses. In any case, these studies should provide an in-depth analysis of the genomic regions analysed, follow more robust approaches, such

as performing discovery and replication analysis, and applying correction for multiple testing. Moreover, although SNPs are a type of genetic variation, the catalog of human genetic variants also includes polymorphic inversions and translocations, tandem repeats, indels and copy number variations, that have been demonstrated to be implicated in asthma (Oliveira et al., 2018; González et al., 2014). Furthermore, future studies should address the gap in the knowledge of the role of non-autosomal variation in AEs. In addition, despite the fact that this doctoral thesis has explored genetic variation in ethnically diverse populations following a multi-ancestry meta-analysis, candidate-gene association studies and specific phenotyping GWAS in minority populations, and admixture mapping in Latinos, further efforts should be made to design and conduct genomic studies of AEs in Asian and African-descent populations. Likewise, although we explored the applicability of PRS in AEs and failed to obtain a good discriminatory PRS, PRS estimation is a quickly evolving field. This together with the increase in sample size of the base and target datasets in future studies may improve the likelihood of finding PRS with good discriminatory capability (Choi et al., 2020).

In terms of admixed individuals, larger sample sizes may be required to detect significant associations due to the inherent complex genetic heterogeneity of these ethnic groups. However, alternative methods incorporating local ancestry into GWAS in admixed populations could increase the statistical power to detect associations (Atkinson et al., 2021). In these populations, the role of local ancestry and interactions of local ancestry with environmental exposures can be an adequate strategy to identify relevant genomic regions that could be subsequently fine-mapped in order to establish the genetic variants that account for the admixture mapping peaks. In this regard, NGS approaches may clarify the role of rare genetic variation, structural, and non-coding variants in AEs. However, these studies are not to be limited only to specific genomic regions of interest, whole-genome sequencing analysis along with integrative genomics strategies could prioritize genetic variants involved in the pathogenesis of AEs based on functional evidence.

6

Conclusions

Based on the results of this thesis, the following conclusions can be drawn:

1. Most genetic studies of AEs performed prior to 2019 focused on candidate genes previously related to asthma susceptibility and/or asthma-related traits, and only two non-pharmacogenetic GWAS of AEs had been performed. The limitations of those studies highlighted the need for more comprehensive studies analyzing diverse populations and assessing the functional consequences of the associations described.

2. A regulatory polymorphism in the *DNASE1L3* gene, related to its gene expression levels and RNA isoform production, was associated with AEs in Hispanics/Latinos and African Americans but not in other populations.

3. A genome-wide significant association of a variant that regulates gene expression levels of *LINC01913* and DNA methylation status of *PKDCC* was uncovered in Hispanics/Latinos and African Americans, but it was not associated with AEs in other populations.

4. Leveraging local ancestry was shown to be helpful in the search of genetic signals for AEs in admixed populations. Genetic polymorphisms at the *STK32A* and *C5orf46* loci were significantly and consistently associated with AEs in Hispanic/Latinos but showed no association in other populations. These loci were associated with *SCGB3A2* gene expression and *DYPSL3* DNA methylation.

5. Two loci at the *VCAM1*/*EXTL2* and *PANK1* regions were associated with AEs and showed consistent effects across multiple populations. The variants identified were associated with DNA methylation and gene expression levels of the genes where they are located.

6. From the 47 variants for AEs or moderate-to-severe-asthma identified by literature mining that were assessed for association with AEs in independent populations, two showed consistent effects at $p<0.05$ in more than two ethnic groups. The association of rs9275503 (*HLA-DQB1*) with AEs, initially described in British adults, was validated in European descents and Hispanics/Latinos. Moreover, the pleiotropic SNP rs72743461 (*SMAD3*), previously associated with moderate-to-severe asthma, was also associated with AEs in admixed populations.

References

Acosta-Herrera, M., Kerick, M., González-Serna, D., Myositis Genetics Consortium, Scleroderma Genetics Consortium, Wijmenga, C., Franke, A., Gregersen, P. K., Padyukov, L., Worthington, J., and et al. (2019). Genome-wide meta-analysis reveals shared new loci in systemic seropositive rheumatic diseases. *Ann Rheum Dis*, 78(3):311–319.

Acosta-Herrera, M., Pino-Yanes, M., Ma, S.-F., Barreto-Luis, A., Corrales, A., Cumplido, J., Pérez-Rodríguez, E., Campo, P., Eng, C., García-Robaina, J. C., and et al. (2015). Fine mapping of the myosin light chain kinase (MYLK) gene replicates the association with asthma in populations of Spanish descent. *J Allergy Clin Immunol*, 136(4):1116–1118.e9.

Akinbami, L. J., Moorman, J. E., Bailey, C., Zahran, H. S., King, M., Johnson, C. A., and Liu, X. (2012). Trends in asthma prevalence, health care use, and mortality in the United States, 2001-2010. *NCHS Data Brief*, 1(94):1–8.

Akinbami, L. J., Moorman, J. E., Simon, A. E., and Schoendorf, K. C. (2014). Trends in racial disparities for asthma outcomes among children 0 to 17 years, 2001-2010. *J Allergy Clin Immunol*, 134(3):547–553.e5.

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*, 19(9):1655–1664.

Ali, Z., Dirks, C. G., and Ulrik, C. S. (2013). Long-term mortality among adults with asthma: a 25-year follow-up of 1,075 outpatients with asthma. *Chest*, 143(6):1649–1655.

Amanat, S., Requena, T., and Lopez-Escamez, J. A. (2020). A systematic review of extreme phenotype strategies to search for rare variants in genetic studies of complex disorders. *Genes*, 11(9):987.

Andiappan, A. K., Anantharaman, R., Nilkanth, P. P., Wang, D. Y., and Chew, F. T. (2010). Evaluating the transferability of Hapmap SNPs to a Singapore Chinese population. *BMC Genet*, 11:36.

Andiappan, A. K., Puan, K. J., Lee, B., Nardin, A., Poidinger, M., Connolly, J., Chew, F. T., Wang, D. Y., and Rotzschke, O. (2014). Allergic airway diseases in a tropical urban environment are driven by dominant mono-specific sensitization against house dust mites. *Allergy*, 69(4):501–509.

Andiappan, A. K., Wang, D. Y., Anantharaman, R., Parate, P. N., Suri, B. K., Low, H. Q., Li, Y., Zhao, W., Castagnoli, P., Liu, J., and et al. (2011). Genome-wide association study for atopy and allergic rhinitis in a Singapore Chinese population. *PloS One*, 6(5):e19719.

Astle, W. J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A. L., Mead, D., Bouman, H., Riveros-Mckay, F., Kostadima, M. A., and et al. (2016). The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell*, 167(5):1415–1429.e19.

Atkinson, E. G., Maihofer, A. X., Kanai, M., Martin, A. R., Karczewski, K. J., Santoro, M. L., Ulirsch, J. C., Kamatani, Y., Okada, Y., Finucane, H. K., and et al. (2021). Tractor uses local ancestry to enable the inclusion of admixed individuals in GWAS and to boost power. *Nat Genet*, 53(2):195–204.

Bai, T. R., Vonk, J. M., Postma, D. S., and Boezen, H. M. (2007). Severe exacerbations predict excess lung function decline in asthma. *Eur Respir J*, 30(3):452–456.

Baines, K. J., Simpson, J. L., Wood, L. G., Scott, R. J., Fibbens, N. L., Powell, H., Cowan, D. C., Taylor, D. R., Cowan, J. O., and Gibson, P. G. (2014). Sputum gene expression signature of 6 biomarkers discriminates asthma inflammatory phenotypes. *J Allergy Clin Immunol*, 133(4):997–1007.

Banno, A., Reddy, A. T., Lakshmi, S. P., and Reddy, R. C. (2018). PPARs: key regulators of airway inflammation and potential therapeutic targets in asthma. *Nucl Receptor Res*, 5:101306.

Barreto, M. L., Cunha, S. S., Alcântara-Neves, N., Carvalho, L. P., Cruz, A. A., Stein, R. T., Genser, B., Cooper, P. J., and Rodrigues, L. C. (2006). Risk factors and immunological pathways for asthma and other allergic diseases in children: background and methodology of a longitudinal study in a large urban center in Northeastern Brazil (Salvador-SCAALA study). *BMC Pulm Med*, 6:15.

Basnet, S., Bochkov, Y. A., Brockman-Schneider, R. A., Kuipers, I., Aesif, S. W., Jackson, D. J., Lemanske, R. F., Ober, C., Palmenberg, A. C., and Gern, J. E. (2019). CDHR3 asthma-risk genotype affects susceptibility of airway epithelium to rhinovirus C infections. *Am J Respir Cell Mol Biol*, 61(4):450–458.

Basu, K., Palmer, C. N. A., Tavendale, R., Lipworth, B. J., and Mukhopadhyay, S. (2009). Adrenergic beta(2)-receptor genotype predisposes to exacerbations in steroid-treated asthmatic patients taking frequent albuterol or salmeterol. *J Allergy Clin Immunol*, 124(6):1188–1194.e3.

Beasley, R., Semprini, A., and Mitchell, E. A. (2015). Risk factors for asthma: is prevention possible? *Lancet*, 386(9998):1075–1085.

Berce, V., Kozmus, C. E. P., and Potočnik, U. (2013). Association among ORMDL3 gene expression, 17q21 polymorphism and response to treatment with inhaled corticosteroids in children with asthma. *Pharmacogenomics J*, 13(6):523–529.

Berthon, B. S., Gibson, P. G., Wood, L. G., MacDonald-Wicks, L. K., and Baines, K. J. (2017). A sputum gene expression signature predicts oral corticosteroid response in asthma. *Eur Respir J*, 49(6):1700180.

Beydon, N., Davis, S. D., Lombardi, E., Allen, J. L., Arets, H. G. M., Aurora, P., Bisgaard, H., Davis, G. M., Ducharme, F. M., Eigen, H., and et al. (2007). An official American Thoracic Society/European Respiratory Society statement: pulmonary function testing in preschool children. *Am J Respir Crit Care Med*, 175(12):1304–1345.

Bin, L.-H., Nielson, L. D., Liu, X., Mason, R. J., and Shu, H.-B. (2003). Identification of uteroglobin-related protein 1 and macrophage scavenger receptor with collagenous structure as a lung-specific ligand-receptor pair. *J Immunol*, 171(2):924–930.

Bisgaard, H., Bønnelykke, K., Sleiman, P. M., Brasholt, M., Chawes, B., Kreiner-Møller, E., Stage, M., Kim, C., Tavendale, R., Baty, F., and et al. (2009). Chromosome 17q21 gene variants are associated with asthma and exacerbations but not atopy in early childhood. *Am J Respir Crit Care Med*, 179(3):179–185.

Black, M. H., Zhou, H., Takayanagi, M., Jacobsen, S. J., and Koebnick, C. (2013). Increased asthma risk and asthma-related health care complications associated with childhood obesity. *Am J Epidemiol*, 178(7):1120–1128.

Blekic, M., Kljaic Bukvic, B., Aberle, N., Marinho, S., Hankinson, J., Custovic, A., and Simpson, A. (2013). 17q12-21 and asthma: interactions with early-life environmental exposures. *Ann Allergy Asthma Immun*, 110(5):347–353.e2.

Bloom, C. I., Palmer, T., Feary, J., Quint, J. K., and Cullinan, P. (2019). Exacerbation patterns in adults with asthma in England. A population-based study. *Am J Respir Crit Care Med*, 199(4):446–453.

Bordoli, M. R., Yum, J., Breitkopf, S. B., Thon, J. N., Italiano, J. E., Xiao, J., Worby, C., Wong, S.-K., Lin, G., Edenius, M., and et al. (2014). A secreted tyrosine kinase acts in the extracellular environment. *Cell*, 158(5):1033–1044.

Borrell, L. N., Nguyen, E. A., Roth, L. A., Oh, S. S., Tcheurekdjian, H., Sen, S., Davis, A., Farber, H. J., Avila, P. C., Brigino-Buenaventura, E., and et al. (2013). Childhood obesity and asthma control in the GALA II and SAGE II studies. *Am J Respir Crit Care Med*, 187(7):697–702.

Bourdin, A., Bjermer, L., Brightling, C., Brusselle, G. G., Chanez, P., Chung, K. F., Custovic, A., Diamant, Z., Diver, S., Djukanovic, R., and et al. (2019). ERS/EAACI statement on severe exacerbations in asthma in adults: facts, priorities and key research questions. *Eur Respir J*, 54(3):1900900.

Boutin, R. C., Petersen, C., Woodward, S. E., Serapio-Palacios, A., Bozorgmehr, T., Loo, R., Chalanuchpong, A., Cirstea, M., Lo, B., Huus, K. E., and et al. (2021). Bacterial–fungal interactions in the neonatal gut influence asthma outcomes later in life. *eLife*, 10:e67740.

Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S., and Davey Smith, G. (2013). Cohort Profile: the 'children of the 90s'– the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol*, 42(1):111–127.

Breeze, C. E., Reynolds, A. P., van Dongen, J., Dunham, I., Lazar, J., Neph, S., Vierstra, J., Bourque, G., Teschendorff, A. E., Stamatoyannopoulos, J. A., and et al. (2019). eFORGE v2.0: updated analysis of cell type-specific signal in epigenomic data. *Bioinformatics*, 35(22):4767–4769.

Brehm, J. M., Acosta-Pérez, E., Klei, L., Roeder, K., Barmada, M. M., Boutaoui, N., Forno, E., Cloutier, M. M., Datta, S., Kelly, R., and et al. (2012). African ancestry and lung function in Puerto Rican children. *J Allergy Clin Immunol*, 129(6):1484–1490.e6.

Briggs, A., Nasser, S., Hammerby, E., Buchs, S., and Virchow, J. C. (2021). The impact of moderate and severe asthma exacerbations on quality of life: a post hoc analysis of randomised controlled trial data. *J Patient Rep Outcomes*, 5(1):6.

British Thoracic Society and Scottish Intercollegiate Guidelines Network (2014). British guideline on the management of asthma. *Thorax*, 69 Suppl 1:1–192.

British Thoracic Society and Scottish Intercollegiate Guidelines Network (2019). British guideline on the management of asthma. SIGN 158. Available from: https://www.brit-thoracic.org.uk/quality-improvement/guidelines/asthma/.

Bronson, P. G., Chang, D., Bhangale, T., Seldin, M. F., Ortmann, W., Ferreira, R. C., Urcelay, E., Pereira, L. F., Martin, J., Plebani, A., and et al. (2016). Common variants at PVT1, ATG13-AMBRA1, AHI1 and CLEC16A are associated with selective IgA deficiency. *Nat Genet*, 48(11):1425–1429.

Browning, S. R. and Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am J Hum Genet*, 81(5):1084–1097.

Bryc, K., Durand, E. Y., Macpherson, J. M., Reich, D., and Mountain, J. L. (2015). The genetic ancestry of African Americans, Latinos, and European Americans across the United States. *Am J Hum Genet*, 96(1):37–53.

Bukvic, B. K., Blekic, M., Simpson, A., Marinho, S., Curtin, J. A., Hankinson, J., Aberle, N., and Custovic, A. (2013). Asthma severity, polymorphisms in 20p13 and their interaction with tobacco smoke exposure. *Pediatr Allergy Immunol*, 24(1):10–18.

Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., and et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*, 47(D1):D1005–D1012.

Busse, W. W. (2017). A role for neutrophils in asthma exacerbations. *Nat Med*, 23(6):658–659.

Bønnelykke, K., Sleiman, P., Nielsen, K., Kreiner-Møller, E., Mercader, J. M., Belgrave, D., Den Dekker, H. T., Husby, A., Sevelsted, A., Faura-Tellez, G., and et al. (2014). A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. *Nat Genet*, 46(1):51–55.

Canals, I., Benetó, N., Cozar, M., Vilageliu, L., and Grinberg, D. (2015). EXTL2 and EXTL3 inhibition with siRNAs as a promising substrate reduction therapy for Sanfilippo C syndrome. *Sci Rep*, 5:13654.

Cardet, J. C. and Casale, T. B. (2019). New insights into the utility of omalizumab. *J Allergy Clin Immunol*, 143(3):923–926.e1.

Centers for Disease Control and Prevention (2018). National surveillance of asthma: United States, 2016-2018. Available from: http://www.cdc.gov/asthma.

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4(1):7.

Chasm, R. M., Pei, Y. V., Pallin, D. J., Shapiro, N., Snyder, B., Sullivan, A. F., Camargo, C. A., and Hasegawa, K. (2015). Sex differences in risk of hospitalization among emergency department patients with acute asthma. *Ann Allergy Asthma Immunol*, 115(1):70–72.e1.

Chen, L. C., Tseng, H. M., Wu, C. J., Kuo, M. L., Wu, C. J., Gao, P. S., Yeh, K. W., Yao, T. C., Lee, W. I., Ou, L. S., and et al. (2012). Evaluation of a common variant of the gene encoding Clara Cell 10 kd Protein (CC10) as a candidate determinant for asthma severity and steroid responsiveness among chinese children. *J Asthma*, 49(7):665–672.

Chiba, Y., Srisodsai, A., Supavilai, P., and Kimura, S. (2005). Interleukin-5 reduces the expression of uteroglobin-related protein (UGRP) 1 gene in allergic airway inflammation. *Immunol Lett*, 97(1):123–129.

Chipps, B. E., Haselkorn, T., Rosén, K., Mink, D. R., Trzaskoma, B. L., and Luskin, A. T. (2018). Asthma exacerbations and triggers in children in TENOR: Impact on quality of life. *J Allergy Clin Immunol Pract*, 6(1):169–176.e2.

Cho, S. H., Min, J.-Y., Kim, D. Y., Oh, S. S., Torgerson, D. R., Pino-Yanes, M., Hu, D., Sen, S., Huntsman, S., Eng, C., and et al. (2016). Association of a PAI-1 gene polymorphism and early life infections with asthma risk, exacerbations, and reduced lung function. *PloS One*, 11(8):e0157848.

Choi, S. W., Mak, T. S.-H., and O'Reilly, P. F. (2020). Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc*, 15(9):2759–2772.

Consortium, . G. P., Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., Marchini, J. L., McCarthy, S., McVean, G. A., and et al. (2015). A global reference for human genetic variation. *Nature*, 526(7571):68–74.

Crapo, O., Hankinson, J. L., Irvin, C., MacIntyre, N. R., Voter, K. Z., Wise, R. A., Graham, B., O'Donnell, C., Paoletti, P., Roca, J., and et al. (1995). Standardization of spirometry: 1994 Update. *Am J Respir Crit Care Med*, 152(3):1107–1136.

Cui, A., Xiang, M., Xu, M., Lu, P., Wang, S., Zou, Y., Qiao, K., Jin, C., Li, Y., Lu, M., and et al. (2019). VCAM-1-mediated neutrophil infiltration exacerbates ambient fine particle-induced lung injury. *Toxicol Lett*, 302:60–74.

Dahlin, A., Denny, J., Roden, D. M., Brilliant, M. H., Ingram, C., Kitchner, T. E., Linneman, J. G., Shaffer, C. M., Weeke, P., Xu, H., and et al. (2015). CMTR1 is associated with increased asthma exacerbations in patients taking inhaled corticosteroids. *Immun Inflamm Dis*, 3(4):350–359.

Dahlin, A., Sordillo, J. E., McGeachie, M., Kelly, R. S., Tantisira, K., Lutz, S. M., Lasky-Su, J., and Wu, A. C. (2020). Genome-wide interaction study reveals age-dependent determinants of responsiveness to inhaled corticosteroids in individuals with asthma. *PloS One*, 15(3):e0229241.

Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A. E., Kwong, A., Vrieze, S. I., Chew, E. Y., Levy, S., McGue, M., and et al. (2016). Next-generation genotype imputation service and methods. *Nat Genet*, 48(10):1284–1287.

Daya, M., Rafaels, N., Brunetti, T. M., Chavan, S., Levin, A. M., Shetty, A., Gignoux, C. R., Boorgula, M. P., Wojcik, G., Campbell, M., and et al. (2019). Association study in African-admixed populations across the Americas recapitulates asthma risk loci in non-African populations. *Nat Commun*, 10(1):880.

Dayem Ullah, A. Z., Oscanoa, J., Wang, J., Nagano, A., Lemoine, N. R., and Chelala, C. (2018). SNPnexus: Assessing the functional relevance of genetic variation to facilitate the promise of precision medicine. *Nucleic Acids Res*, 46(W1):W109–W113.

de Roos, E. W., Lahousse, L., Verhamme, K. M., Braunstahl, G.-J., in 't Veen, J. C., Stricker, B. H., and Brusselle, G. G. (2021). Incidence and predictors of asthma exacerbations in middle-aged and older adults: the Rotterdam Study. *ERJ Open Res*, 7(3):00126–02021.

De Simoni, A., Horne, R., Fleming, L., Bush, A., and Griffiths, C. (2017). What do adolescents with asthma really think about adherence to inhalers? Insights from a qualitative analysis of a UK online forum. *BMJ Open*, 7(6):e015245.

Deem, T. L. and Cook-Mills, J. M. (2004). Vascular cell adhesion molecule 1 (VCAM-1) activation of endothelial cell matrix metalloproteinases: role of reactive oxygen species. *Blood*, 104(8):2385–2393.

Del Valle-Pérez, B., Martínez, V. G., Lacasa-Salavert, C., Figueras, A., Shapiro, S. S., Takafuta, T., Casanovas, O., Capellà, G., Ventura, F., and Viñals, F. (2010). Filamin B plays a key role in vascular endothelial growth factor-induced endothelial cell motility through its interaction with Rac-1 and Vav-2. *J Biol Chem*, 285(14):10748–10760.

Denlinger, L. C., Manthei, D. M., Seibold, M. A., Ahn, K., Bleecker, E., Boushey, H. A., Calhoun, W. J., Castro, M., Chinchili, V. M., Fahy, J. V., and et al. (2013). P2X7-regulated protection from exacerbations and loss of control is independent of asthma maintenance therapy. *Am J Respir Crit Care Med*, 187(1):28–33.

Denlinger, L. C., Phillips, B. R., Ramratnam, S., Ross, K., Bhakta, N. R., Cardet, J. C., Castro, M., Peters, S. P., Phipatanakul, W., Aujla, S., and et al. (2017). Inflammatory and comorbid features of patients with severe asthma and frequent exacerbations. *Am J Respir Crit Care Med*, 195(3):302–313.

Dharmage, S. C., Perret, J. L., and Custovic, A. (2019). Epidemiology of asthma in children and adults. *Front Pediatr*, 7:246.

di Palmo, E., Cantarelli, E., Catelli, A., Ricci, G., Gallucci, M., Miniaci, A., and Pession, A. (2021). The predictive role of biomarkers and genetics in childhood asthma exacerbations. *Int J Mol Sci*, 22(9):4651.

Ding, L., Ren, J., Zhang, D., Li, Y., Huang, X., Hu, Q., Wang, H., Song, Y., Ni, Y., and Hou, Y. (2018). A novel stromal lncRNA signature reprograms fibroblasts to promote the growth of oral squamous cell carcinoma via LncRNA-CAF / interleukin-33. *Carcinogenesis*, 39(3):397–406.

Ding, Y., Yang, H., He, H., Shi, H., He, P., Yan, S., and Jin, T. (2015). Plasma adiponectin concentrations and adiponectin gene polymorphisms are associated with bronchial asthma in the Chinese Li population. *Iran J Allergy Asthma Immunol*, 14(3):292–297.

Du, P., Zhang, X., Huang, C.-C., Jafari, N., Kibbe, W. A., Hou, L., and Lin, S. M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, 11:587.

Du, R., Litonjua, A. A., Tantisira, K. G., Lasky-Su, J., Sunyaev, S. R., Klanderman, B. J., Celedón, J. C., Avila, L., Soto-Quiros, M. E., and Weiss, S. T. (2012). Genome-wide association study reveals class I MHC-restricted T cell-associated molecule gene (CRTAM) variants interact with vitamin D levels to affect asthma exacerbations. *J Allergy Clin Immunol*, 129(2):368–373, 373.e1–5.

Duncan, L. E., Ostacher, M., and Ballon, J. (2019). How genome-wide association studies (GWAS) made traditional candidate gene studies obsolete. *Neuropsychopharmacology*, 44(9):1518–1523.

D'Amato, G., Vitale, C., Molino, A., Stanziola, A., Sanduzzi, A., Vatrella, A., Mormile, M., Lanza, M., Calabrese, G., Antonicelli, L., and et al. (2016). Asthma-related deaths. *Multidiscip Respir Med*, 11:37.

Enaud, R., Prevel, R., Ciarlo, E., Beaufils, F., Wieërs, G., Guery, B., and Delhaes, L. (2020). The gut-lung axis in health and respiratory diseases: A place for inter-organ and inter-kingdom crosstalks. *Front Cell Infect Microbiol*, 10:9.

Engavale, M., McCord, J., Mapp, B., Nzimulinda, N., Bengtson, E., Sutton, R. B., and Keyel, P. A. (2021). Dnase1 family in autoimmunity. *Encyclopedia*, 1(3):527–541.

Engelkes, M., Baan, E. J., Ridder, M. A. J., Svensson, E., Prieto-Alhambra, D., Lapi, F., Giaquinto, C., Picelli, G., Boudiaf, N., Albers, F., and et al. (2020a). Incidence, risk factors and re-exacerbation rate of severe asthma exacerbations in a multinational, multi-database pediatric cohort study. *Pediatr Allergy Immunol*, 31(5):496–505.

Engelkes, M., de Ridder, M. A., Svensson, E., Berencsi, K., Prieto-Alhambra, D., Lapi, F., Giaquinto, C., Picelli, G., Boudiaf, N., Albers, F. C., and et al. (2020b). Multinational cohort study of mortality in patients with asthma and severe asthma. *Respir Med*, 165:105919.

Engelkes, M., Janssens, H. M., de Jongste, J. C., Sturkenboom, M. C., and Verhamme, K. M. (2015). Medication adherence and the risk of severe asthma exacerbations: a systematic review. *Eur Respir J*, 45(2):396–407.

Erbas, B., Jazayeri, M., Lambert, K. A., Katelaris, C. H., Prendergast, L. A., Tham, R., Parrodi, M. J., Davies, J., Newbigin, E., Abramson, M. J., and et al. (2018). Outdoor pollen is a trigger of child and adolescent asthma emergency department presentations: A systematic review and meta-analysis. *Allergy*, 73(8):1632–1641.

Esmaeilzadeh, H., Nouri, F., Nabavizadeh, S. H., Alyasin, S., and Mortazavi, N. (2021). Can eosinophilia and neutrophil-lymphocyte ratio predict hospitalization in asthma exacerbation? *Allergy Asthma Clin Immunol*, 17(1):16.

Everman, J. L., Sajuthi, S., Saef, B., Rios, C., Stoner, A. M., Numata, M., Hu, D., Eng, C., Oh, S., Rodriguez-Santana, J., and et al. (2019). Functional genomics of CDHR3 confirms its role in HRV-C infection and childhood asthma exacerbations. *J Allergy Clin Immunol*, 144(4):962–971.

Eyre, S., Bowes, J., Diogo, D., Lee, A., Barton, A., Martin, P., Zhernakova, A., Stahl, E., Viatte, S., McAllister, K., and et al. (2012). High-density genetic mapping identifies new susceptibility loci for rheumatoid arthritis. *Nat Genet*, 44(12):1336–1340.

Fallon, M. E. and Hinds, M. T. (2021). Single cell morphological metrics and cytoskeletal alignment regulate VCAM-1 protein expression. *Biochem Biophys Res Commun*, 555:160–167.

Farhan, H. (2020). Tyrosine kinase signaling in and on the endoplasmic reticulum. *Biochem Soc Trans*, 48(1):199–205.

Fazlollahi, M., Lee, T. D., Andrade, J., Oguntuyo, K., Chun, Y., Grishina, G., Grishin, A., and Bunyavanich, S. (2018). The nasal microbiome in asthma. *J Allergy Clin Immunol*, 142(3):834–843.e2.

Fehrenbach, H., Wagner, C., and Wegmann, M. (2017). Airway remodeling in asthma: what really matters. *Cell Tissue Res*, 367(3):551–569.

Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., and et al. (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)*, 2017.

Fitzpatrick, A. M., Chipps, B. E., Holguin, F., and Woodruff, P. G. (2020). T2-"Low" asthma: Overview and management strategies. *J Allergy Clin Immunol Pract*, 8(2):452–463.

Flores, C., Ma, S.-F., Pino-Yanes, M., Wade, M. S., Pérez-Méndez, L., Kittles, R. A., Wang, D., Papaiahgari, S., Ford, J. G., Kumar, R., and et al. (2012). African ancestry is associated with asthma risk in African Americans. *PloS One*, 7(1):e26807.

Fonseca, H., da Silva, T. M., Saraiva, M., Santolalla, M. L., Sant'Anna, H. P., Araujo, N. M., Lima, N. P., Rios, R., Tarazona-Santos, E., Horta, B. L., and et al. (2020). Genomic regions 10q22.2, 17q21.31, and 2p23.1 can contribute to a lower lung function in African descent populations. *Genes*, 11(9):1047.

Forno, E., Lescher, R., Strunk, R., Weiss, S., Fuhlbrigge, A., Celedón, J. C., and Group, C. A. M. P. R. (2011). Decreased response to inhaled steroids in overweight and obese asthmatic children. *J Allergy Clin Immunol*, 127(3):741–749.

Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., Henderson, J., Macleod, J., Molloy, L., Ness, A., and et al. (2013). Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol*, 42(1):97–110.

Frey, A., Lunding, L. P., Ehlers, J. C., Weckmann, M., Zissler, U. M., and Wegmann, M. (2020). More than just a barrier: The immune functions of the airway epithelium in asthma pathogenesis. *Front Immunol*, 11:761.

Fricker, M., Gibson, P. G., Powell, H., Simpson, J. L., Yang, I. A., Upham, J. W., Reynolds, P. N., Hodge, S., James, A. L., Jenkins, C., Peters, M. J., Marks, G. B., Baraket, M., and Baines, K. J. (2019). A sputum 6-gene signature predicts future exacerbations of poorly controlled asthma. *J Allergy Clin Immunol*, 144(1):51–60.e11.

Frøssing, L., Silberbrandt, A., Bülow, A. V., Klein, D. K., Christensen, M. R., Backer, V., Baines, K. J., and Porsbjerg, C. (2021). Airway gene expression identifies subtypes of type 2 inflammation in severe asthma. *Clin Exp Allergy*, 52(1):59–69.

Fuchs, O., Bahmer, T., Weckmann, M., Dittrich, A.-M., Schaub, B., Rösler, B., Happle, C., Brinkmann, F., Ricklefs, I., König, I. R., and et al. (2018). The all age asthma cohort (ALLIANCE) - from early beginnings to chronic disease: a longitudinal cohort study. *BMC Pulm Med*, 18(1):140.

Galanter, J. M., Gignoux, C. R., Torgerson, D. G., Roth, L. A., Eng, C., Oh, S. S., Nguyen, E. A., Drake, K. A., Huntsman, S., Hu, D., and et al. (2014). Genome-wide association study and admixture mapping identify different asthma-associated loci in Latinos: the Genes-environments & Admixture in Latino Americans study. *J Allergy Clin Immunol*, 134(2):295–305.

GBD 2015 Chronic Respiratory Disease Collaborators (2017). Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Respir Med*, 5(9):691–706.

Ghoussaini, M., Mountjoy, E., Carmona, M., Peat, G., Schmidt, E. M., Hercules, A., Fumis, L., Miranda, A., Carvalho-Silva, D., Buniello, A., and et al. (2021). Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. *Nucleic Acids Res*, 49(D1):D1311–D1320.

Gibson, P. G., Yang, I. A., Upham, J. W., Reynolds, P. N., Hodge, S., James, A. L., Jenkins, C., Peters, M. J., Marks, G. B., Baraket, M., and et al. (2019). Efficacy of azithromycin in severe asthma from the AMAZES randomised trial. *ERJ Open Res*, 5(4):00056–2019.

Gignoux, C. R., Torgerson, D. G., Pino-Yanes, M., Uricchio, L. H., Galanter, J., Roth, L. A., Eng, C., Hu, D., Nguyen, E. A., Huntsman, S., and et al. (2019). An admixture mapping meta-analysis implicates genetic variation at 18q21 with asthma susceptibility in Latinos. *J Allergy Clin Immunol*, 143(3):957–969.

Girouard, M.-P., Simas, T., Hua, L., Morquette, B., Khazaei, M. R., Unsain, N., Johnstone, A. D., Rambaldi, I., Sanz, R. L., Di Raddo, M.-E., and et al. (2020). Collapsin response mediator protein 4 (CRMP4) facilitates wallerian degeneration and axon regeneration following sciatic nerve injury. *eNeuro*, 7(2):ENEURO.0479–19.2020.

Global Initiative for Asthma (2021a). Global strategy for asthma management and prevention. Available from: http://ginasthma.org/.

Global Initiative for Asthma (2021b). Global strategy for asthma management and prevention. Archived reports. Available from: https://ginasthma.org/archived-reports/.

González, J. R., Cáceres, A., Esko, T., Cuscó, I., Puig, M., Esnaola, M., Reina, J., Siroux, V., Bouzigon, E., Nadif, R., and et al. (2014). A common 16p11.2 inversion underlies the joint susceptibility to asthma and obesity. *Am J Hum Genet*, 94(3):361–372.

Grainge, C. L., Lau, L. C., Ward, J. A., Dulay, V., Lahiff, G., Wilson, S., Holgate, S., Davies, D. E., and Howarth, P. H. (2011). Effect of bronchoconstriction on airway remodeling in asthma. *N Engl J Med*, 364(21):2006–2015.

Grossman, N. L., Ortega, V. E., King, T. S., Bleecker, E. R., Ampleford, E. A., Bacharier, L. B., Cabana, M. D., Cardet, J. C., Carr, T. F., Castro, M., and et al. (2019). Exacerbation-prone asthma in the context of race and ancestry in Asthma Clinical Research Network trials. *J Allergy Clin Immunol*, 144(6):1524–1533.

Gruffydd-Jones, K., Thomas, M., Roman-Rodríguez, M., Infantino, A., Fitzgerald, J. M., Pavord, I., Haddon, J. M., Elsasser, U., and Vogelberg, C. (2019). Asthma impacts on workplace productivity in employed patients who are symptomatic despite background therapy: A multinational survey. *J Asthma Allergy*, 12:183–194.

Grunwell, J. R., Travers, C., and Fitzpatrick, A. M. (2018). Inflammatory and comorbid features of children admitted to a PICU for status asthmaticus. *Pediatr Crit Care Med*, 19(11):e585–e594.

GTEx Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*, 369(6509):1318–1330.

Guillen-Guio, B., Hernández-Beeftink, T., Marcelino-Rodriguez, I., Rodríguez-Pérez, H., Lorenzo-Salazar, J. M., Espinilla-Peña, M., Corrales, A., Pino-Yanes, M., Callero, A., Perez-Rodriguez, E., and et al. (2020). Admixture mapping of asthma in southwestern Europeans with North African ancestry influences. *Am J Physiol Lung Cell Mol Physiol*, 318(5):L965–L975.

Guxens, M., Ballester, F., Espada, M., Fernández, M. F., Grimalt, J. O., Ibarluzea, J., Olea, N., Rebagliato, M., Tardón, A., Torrent, M., and et al. (2012). Cohort Profile: the INMA–INfancia y Medio Ambiente–(Environment and Childhood) Project. *Int J Epidemiol*, 41(4):930–940.

GWAS Diversity Monitor (2020). GWAS Diversity Monitor. Available from: https://gwasdiversitymonitor.com.

Haag, K., Fraser, A., Hiller, R., Seedat, S., Zimmerman, A., and Halligan, S. L. (2020). The emergence of sex differences in PTSD symptoms across development: evidence from the ALSPAC cohort. *Psychol Med*, 50(10):1755–1760.

Hallas, H. W., Chawes, B. L., Arianto, L., Rasmussen, M. A., Kunøe, A., Stokholm, J., Bønnelykke, K., and Bisgaard, H. (2020). Children with asthma have fixed airway obstruction through childhood unaffected by exacerbations. *J Allergy Clin Immunol Pract*, 8(4):1263–1271.e3.

Hammond, R. K., Pahl, M. C., Su, C., Cousminer, D. L., Leonard, M. E., Lu, S., Doege, C. A., Wagley, Y., Hodge, K. M., Lasconi, C., and et al. (2021). Biological constraints on GWAS SNPs at suggestive significance thresholds reveal additional BMI loci. *eLife*, 10.

Han, B. and Eskin, E. (2011). Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am J Hum Genet*, 88(5):586–598.

Hansen, K. D. (2016). Illuminahumanmethylation450kanno.ilmn12.hg19: Annotation for illumina's 450k methylation arrays. r package version 0.6.0.

Harada, H., Farhani, N., Wang, X.-F., Sugita, S., Charish, J., Attisano, L., Moran, M., Cloutier, J.-F., Reber, M., Bremner, R., and et al. (2019). Extracellular phosphorylation drives the formation of neuronal circuitry. *Nat Chem Biol*, 15(11):1035–1042.

Hardin, M., Cho, M. H., McDonald, M. L., Wan, E., Lomas, D. A., Coxson, H. O., MacNee, W., Vestbo, J., Yates, J. C., Agusti, A., and et al. (2016). A genome-wide analysis of the response to inhaled β 2-agonists in chronic obstructive pulmonary disease. *Pharmacogenomics J*, 16(4):326–335.

Hasegawa, K., Calhoun, W. J., Pei, Y. V., Chasm, R. M., Youngquist, S. T., Bittner, J. C., and Camargo, C. A. (2015). Sex differences in hospital length of stay in children and adults hospitalized for asthma exacerbation. *Ann Allergy Asthma Immunol*, 115(6):533–535.e1.

Hawcutt, D. B., Francis, B., Carr, D. F., Jorgensen, A. L., Yin, P., Wallin, N., O'Hara, N., Zhang, E. J., Bloch, K. M., Ganguli, A., and et al. (2018). Susceptibility to corticosteroid-induced adrenal suppression: a genome-wide association study. *Lancet Respir Med*, 6(6):442–450.

Hawcutt, D. B., Jorgensen, A. L., Wallin, N., Thompson, B., Peak, M., Lacy, D., Newland, P., Didi, M., Couriel, J., Blair, J., and et al. (2015). Adrenal responses to a low-dose short synacthen test in children with asthma. *Clin Endocrinol*, 82(5):648–656.

Henneberger, P. K., Mirabelli, M. C., Kogevinas, M., Anto, J. M., Plana, E., Dahlman-Hoglund, A., Jarvis, D. L., Kromhout, H., Lillienberg, L., Norback, D., and et al. (2010). The occupational contribution to severe exacerbation of asthma. *Eur Respir J*, 36(4):743–750.

Hennessy, E. J., Sheedy, F. J., Santamaria, D., Barbacid, M., and O'Neill, L. A. J. (2011). Toll-like receptor-4 (TLR4) down-regulates microRNA-107, increasing macrophage adhesion via cyclin-dependent kinase 6. *J Biol Chem*, 286(29):25531–25539.

Hernandez-Pacheco, N., Farzan, N., Francis, B., Karimi, L., Repnik, K., Vijverberg, S. J., Soares, P., Schieck, M., Gorenjak, M., Forno, E., and et al. (2019). Genome-wide association study of inhaled corticosteroid response in admixed children with asthma. *Clin Exp Allergy*, 49(6):789–798.

Hernandez-Pacheco, N., Flores, C., Oh, S. S., Burchard, E. G., and Pino-Yanes, M. (2016). What ancestry can tell us about the genetic origins of inter-ethnic differences in asthma expression. *Curr Allergy Asthma Rep*, 16(8):53.

Hernandez-Pacheco, N., Vijverberg, S. J., Herrera-Luis, E., Li, J., Sio, Y. Y., Granell, R., Corrales, A., Maroteau, C., Lethem, R., Perez-Garcia, J., and et al. (2021). Genome-wide association study of asthma exacerbations despite inhaled corticosteroid use. *Eur Respir J*, 57(5):2003388.

Hernandez-Pacheco, N., Gorenjak, M., Jurgec, S., Corrales, A., Jorgensen, A., Karimi, L., Vijverberg, S. J., Berce, V., Schieck, M., Acosta-Herrera, M., and et al. (2021). Combined analysis of transcriptomic and genetic data for the identification of loci involved in glucocorticosteroid response in asthma. *Allergy*, 76(4):1238–1243.

Herrera-Luis, E., Hernandez-Pacheco, N., Vijverberg, S. J., Flores, C., and Pino-Yanes, M. (2019). Role of genomics in asthma exacerbations. *Curr Opin Pulm Med*, 25(1):101–112.

Hershey, G. K. (2003). IL-13 receptors and signaling pathways: An evolving web. *J Allergy Clin Immunol*, 111(4):677–690.

Hirai, K., Shirai, T., Suzuki, M., Shimomura, T., and Itoh, K. (2018). Association between (CCTTT)n repeat polymorphism in NOS2 promoter and asthma exacerbations. *J Allergy Clin Immunol*, 142(2):663–665.e3.

Hodkinson, A., Bower, P., Grigoroglou, C., Zghebi, S. S., Pinnock, H., Kontopantelis, E., and Panagioti, M. (2020). Self-management interventions to reduce healthcare use and improve quality of life among patients with asthma: systematic review and network meta-analysis. *BMJ*, 370:m2521.

Hortelano, S., López-Fontal, R., Través, P. G., Villa, N., Grashoff, C., Boscá, L., and Luque, A. (2010). ILK mediates LPS-induced vascular adhesion receptor expression and subsequent leucocyte trans-endothelial migration. *Cardiovasc Res*, 86(2):283–292.

Hough, K. P., Curtiss, M. L., Blain, T. J., Liu, R.-M., Trevor, J., Deshane, J. S., and Thannickal, V. J. (2020). Airway remodeling in asthma. *Front Med*, 7:191.

Howie, B. N., Donnelly, P., and Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*, 5(6):e1000529.

Hunninghake, G. M., Soto-Quirós, M. E., Avila, L., Su, J., Murphy, A., Demeo, D. L., Ly, N. P., Liang, C., Sylvia, J. S., Klanderman, B. J., and et al. (2007). Polymorphisms in IL13, total IgE, eosinophilia, and asthma exacerbations in childhood. *J Allergy Clin Immunol*, 120(1):84–90.

Hunninghake, G. M., Soto-Quirós, M. E., Lasky-Su, J., Avila, L., Ly, N. P., Liang, C., Klanderman, B. J., Raby, B. A., Gold, D. R., Weiss, S. T., and et al. (2008). Dust mite exposure modifies the effect of functional IL10 polymorphisms on allergy and asthma exacerbations. *J Allergy Clin Immunol*, 122(1):93–98.e5.

Ikram, M. A., Brusselle, G. G., Murad, S. D., van Duijn, C. M., Franco, O. H., Goedegebure, A., Klaver, C. C., Nijsten, T. E., Peeters, R. P., Stricker, B. H., and et al. (2017). The Rotterdam Study: 2018 update on objectives, design and main results. *Eur J Epidemiol*, 32(9):807–850.

Ioachimescu, O. C. and Desai, N. S. (2019). Nonallergic triggers and comorbidities in asthma exacerbations and disease severity. *Clin Chest Med*, 40(1):71–85.

Ivanova, J. I., Bergman, R., Birnbaum, H. G., Colice, G. L., Silverman, R. A., and McLaurin, K. (2012). Effect of asthma exacerbations on health care costs among asthmatic patients with moderate and severe persistent asthma. *J Allergy Clin Immunol*, 129(5):1229–1235.

Jayaram, L. (2006). Determining asthma treatment by monitoring sputum cell counts: effect on exacerbations. *Eur Respir J*, 27(3):483–494.

Johansson, Å., Rask-andersen, M., Karlsson, T., and Ek, W. E. (2019). Genome-wide association analysis of 350 000 Caucasians from the UK Biobank identifies novel loci for asthma, hay fever and eczema. *Hum Mol Genet*, 28(23):4022–4041.

Juniper, E. F., O'Byrne, P. M., Guyatt, G. H., Ferrie, P. J., and King, D. R. (1999). Development and validation of a questionnaire to measure asthma control. *Eur Respir J*, 14(4):902–907.

Kaidonis, X., Liaw, W. C., Roberts, A. D., Ly, M., Anson, D., and Byers, S. (2010). Gene silencing of EXTL2 and EXTL3 as a substrate deprivation therapy for heparan sulphate storing mucopolysaccharidoses. *Eur J Hum Genet*, 18(2):194–199.

Kalotas, J. O., Wang, C. J., Noble, P. B., and Wang, K. C. W. (2021). Intrauterine growth restriction promotes postnatal airway hyperresponsiveness independent of allergic disease. *Front Med*, 8:674324.

Kamat, M. A., Blackshaw, J. A., Young, R., Surendran, P., Burgess, S., Danesh, J., Butterworth, A. S., and Staley, J. R. (2019). PhenoScanner v2: an expanded tool for searching human genotype–phenotype associations. *Bioinformatics*, 35(22):4851–4853.

Kanai, M., Tanaka, T., and Okada, Y. (2016). Empirical estimation of genome-wide significance thresholds based on the 1000 Genomes Project data set. *J Hum Genet*, 61(10):861–866.

Kang, H. M. (2016). EPACTS (Efficient and Parallelizable Association Container Toolbox). Available from: https://genome.sph.umich.edu/wiki/EPACTS.

Kang, H. R., Song, H. J., Nam, J. H., Hong, S., Yang, S., Ju, S., Lee, S. W., Kim, T. B., Kim, H., and Lee, E. K. (2018). Risk factors of asthma exacerbation based on asthma severity: a nationwide population-based observational study in South Korea. *BMJ Open*, 8(3):e020825.

Karolchik, D., Hinrichs, A. S., Furey, T. S., Roskin, K. M., Sugnet, C. W., Haussler, D., and Kent, W. J. (2004). The UCSC Table Browser data retrieval tool. *Nucleic Acids Res*, 32(Database issue):D493–D496.

Katta, K., Imran, T., Busse-Wicher, M., Grønning, M., Czajkowski, S., and Kusche-Gullberg, M. (2015). Reduced expression of EXTL2, a member of the Exostosin (EXT) family of glycosyltransferases, in human embryonic kidney 293 cells results in longer heparan sulfate chains. *J Biol Chem*, 290(21):13168–13177.

Kelly, M. M., O'Connor, T. M., Leigh, R., Otis, J., Gwozd, C., Gauvreau, G. M., Gauldie, J., and O'Byrne, P. M. (2010). Effects of budesonide and formoterol on allergen-induced airway responses, inflammation, and airway remodeling in asthma. *J Allergy Clin Immunol*, 125(2):349–356.e13.

Kenyon, C. C., Strane, D., Floyd, G. C., Jacobi, E. G., Penrose, T. J., Ewig, J. M., DaVeiga, S. P., Zorc, J. J., Rubin, D. M., and Bryant-Stephens, T. C. (2020). An asthma population health improvement initiative for children with frequent hospitalizations. *Pediatrics*, 146(5):e20193108.

Keys, K. L., Mak, A. C. Y., White, M. J., Eckalbar, W. L., Dahl, A. W., Mefford, J., Mikhaylova, A. V., Contreras, M. G., Elhawary, J. R., Eng, C., and et al. (2020). On the cross-population generalizability of gene expression prediction models. *PLoS Genet*, 16(8):e1008927.

Kim, J. M., Han, H., Bahn, M., Hur, Y., Yeo, C.-Y., and Kim, D.-W. (2020). Secreted tyrosine kinase Vlk negatively regulates Hedgehog signaling by inducing lysosomal degradation of Smoothened. *Biochem J*, 477(1):121–136.

Kitagawa, H., Shimakawa, H., and Sugahara, K. (1999). The tumor suppressor EXT-like gene EXTL2 encodes an alpha1, 4-N-acetylhexosaminyltransferase that transfers N-acetylgalactosamine and N-acetylglucosamine to the common glycosaminoglycan-protein linkage region. The key enzyme for the chain initiation of heparan sulfate. *J Biol Chem*, 274(20):13933–13937.

Kljaic-Bukvic, B., Blekic, M., Aberle, N., Curtin, J. A., Hankinson, J., Semic-Jusufagic, A., Belgrave, D., Simpson, A., and Custovic, A. (2014). Genetic variants in endotoxin signalling pathway, domestic endotoxin exposure and asthma exacerbations. *Pediatr Allergy Immunol*, 25(6):552–557.

Kochis, S. R., Grant, T., Phipatanakul, W., Perzanowski, M., Balcer-Whaley, S., Curtin-Brosnan, J., Newman, M., Cunningham, A., Bollinger, M. E., and Matsui, E. C. (2019). Caregiver-reported asthma control predicts future visits, independent of guideline-based control measures. *J Allergy Clin Immunol Pract*, 7(5):1516–1521.e1.

Kong, D.-H., Kim, Y., Kim, M., Jang, J., and Lee, S. (2018). Emerging roles of vascular cell adhesion molecule-1 (VCAM-1) in immunological disorders and cancer. *Int J Mol Sci*, 19(4):1057.

Koster, E. S., Maitland-van der Zee, A.-H., Tavendale, R., Mukhopadhyay, S., Vijverberg, S. J. H., Raaijmakers, J. A. M., and Palmer, C. N. A. (2011). FCER2 T2206C variant associated with chronic symptoms and exacerbations in steroid-treated asthmatic children. *Allergy*, 66(12):1546–1552.

Koster, E. S., Raaijmakers, J. A., Koppelman, G. H., Postma, D. S., van der Ent, C. K., Koenderman, L., Bracke, M., and Maitland-van der Zee, A. H. (2009). Pharmacogenetics of anti-inflammatory treatment in children with asthma: rationale and design of the PACMAN cohort. *Pharmacogenomics*, 10(8):1351–1361.

Koyama, R., Arai, T., Kijima, M., Sato, S., Miura, S., Yuasa, M., Kitamura, D., and Mizuta, R. (2016). DNase γ, DNase I and caspase-activated DNase cooperate to degrade dead cells. *Genes Cells*, 21(11):1150–1163.

Krings, J. G., Goss, C. W., Lew, D., Samant, M., McGregor, M. C., Boomer, J., Bacharier, L. B., Sheshadri, A., Hall, C., Brownell, J., and et al. (2021). Quantitative CT metrics are associated with longitudinal lung function decline and future asthma exacerbations: Results from SARP-3. *J Allergy Clin Immunol*, 148(3):752–762.

Krishnan, V., Diette, G. B., Rand, C. S., Bilderback, A. L., Merriman, B., Hansel, N. N., and Krishnan, J. A. (2006). Mortality in patients hospitalized for asthma exacerbations in the United States. *Am J Respir Crit Care Med*, 174(6):633–638.

Ku, C. S., Loy, E. Y., Salim, A., Pawitan, Y., and Chia, K. S. (2010). The discovery of human genetic variations and their use as disease markers: past, present and future. *J Hum Genet*, 55(7):403–415.

Kull, I., Melen, E., Alm, J., Hallberg, J., Svartengren, M., van Hage, M., Pershagen, G., Wickman, M., and Bergström, A. (2010). Breast-feeding in relation to asthma, lung function, and sensitization in young schoolchildren. *J Allergy Clin Immunol*, 125(5):1013–1019.

Kumar, R., Williams, L. K., Kato, A., Peterson, E. L., Favoreto, S., Hulse, K., Wang, D., Beckman, K., Thyne, S., LeNoir, M., and et al. (2012). Genetic variation in B cell-activating factor of the TNF family (BAFF) and asthma exacerbations among African American subjects. *J Allergy Clin Immunol*, 130(4):996–999.e6.

Kuna, P., Peters, M. J., Manjra, A. I., Jorup, C., Naya, I. P., Martínez-Jimenez, N. E., and Buhl, R. (2007). Effect of budesonide/formoterol maintenance and reliever therapy on asthma exacerbations. *Int J Clin Pract*, 61(5):725–736.

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M. J., and et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, 518(7539):317–330.

Kurotani, R., Okumura, S., Matsubara, T., Yokoyama, U., Buckley, J. R., Tomita, T., Kezuka, K., Nagano, T., Esposito, D., Taylor, T. E., and et al. (2011). Secretoglobin 3A2 suppresses bleomycin-induced pulmonary fibrosis by Transforming Growth Factor β signaling down-regulation. *J Biol Chem*, 286(22):19682–19692.

Kynyk, J. A., Mastronarde, J. G., and McCallister, J. W. (2011). Asthma, the sex difference. *Curr Opin Pulm Med*, 17(1):6–11.

Lam, R. W., Inselman, J. W., Jeffery, M. M., Maddux, J. T., and Rank, M. A. (2021). National decline in asthma exacerbations in US during COVID19 pandemic. *Ann Allergy Asthma Immunol*, 127(6):692–694.

Lane, S., Molina, J., and Plusa, T. (2006). An international observational prospective study to determine the Cost Of Asthma eXacerbations (COAX). *Respir Med*, 100(3):434–450.

Lang, J. E., Tang, M., Zhao, C., Hurst, J., Wu, A., and Goldstein, B. A. (2020). Well-child care attendance and risk of asthma exacerbations. *Pediatrics*, 146(6):e20201023.

Lappalainen, U., Whitsett, J. A., Wert, S. E., Tichelaar, J. W., and Bry, K. (2005). Interleukin-1β causes pulmonary inflammation, emphysema, and airway remodeling in the adult murine lung. *Am J Respir Cell Mol Biol*, 32(4):311–318.

Lee, D. S., Gross, E., Hotz, A., and Rastogi, D. (2020a). Comparison of severity of asthma hospitalization between African American and Hispanic children in the Bronx. *J Asthma*, 57(7):736–742.

Lee, E. Y., Mak, A. C. Y., Hu, D., Sajuthi, S., White, M. J., Keys, K. L., Eckalbar, W., Bonser, L., Huntsman, S., Urbanek, C., and et al. (2020b). Whole-genome sequencing identifies novel functional loci associated with lung function in Puerto Rican youth. *Am J Respir Crit Care Med*, 202(7):962–972.

Lee, J. J. and Rosenberg, H. F., editors (2013). *Eosinophil Trafficking*, page 121–166. Academic Press, Boston.

Leiter, K., Franks, K., Borland, M. L., Coleman, L., Harris, L., le Souëf, P. N., and Laing, I. A. (2020). Vitamin D receptor polymorphisms are associated with severity of wheezing illnesses and asthma exacerbations in children. *J Steroid Biochem Mol Biol*, 201:105692.

Lima, J. J., Holbrook, J. T., Wang, J., Sylvester, J. E., Blake, K. V., Blumenthal, M. N., Castro, M., Hanania, N., and Wise, R. (2006a). The C523A beta2 adrenergic receptor polymorphism associates with markers of asthma severity in African Americans. *J Asthma*, 43(3):185–191.

Lima, J. J., Zhang, S., Grant, A., Shao, L., Tantisira, K. G., Allayee, H., Wang, J., Sylvester, J., Holbrook, J., Wise, R., and et al. (2006b). Influence of leukotriene pathway polymorphisms on response to montelukast in asthma. *Am J Respir Crit Care Med*, 173(4):379–385.

Lin, B. M., Grinde, K. E., Brody, J. A., Breeze, C. E., Raffield, L. M., Mychaleckyj, J. C., Thornton, T. A., Perry, J. A., Baier, L. J., de las Fuentes, L., and et al. (2021). Whole genome sequence analyses of eGFR in 23,732 people representing multiple ancestries in the NHLBI trans-omics for precision medicine (TOPMed) consortium. *EBioMedicine*, 63:103157.

Liu, A. H., Zeiger, R., Sorkness, C., Mahr, T., Ostrom, N., Burgess, S., Rosenzweig, J. C., and Manjunath, R. (2007). Development and cross-sectional validation of the Childhood Asthma Control Test. *J Allergy Clin Immunol*, 119(4):817–825.

Liu, H.-Y., Li, C.-X., Liang, Z.-Y., Zhang, S.-Y., Yang, W.-Y., Ye, Y.-M., Lin, Y.-X., Chen, R.-C., Zhou, H.-W., and Su, J. (2020). The interactions of airway bacterial and fungal communities in clinically stable asthma. *Front Microbiol*, 11:1647.

Lloyd, A., Price, D., and Brown, R. (2007). The impact of asthma exacerbations on health-related quality of life in moderate to severe asthma patients in the UK. *Prim Care Respir J*, 16(1):22–27.

Longo, C., Bartlett, G., Schuster, T., Ducharme, F. M., MacGibbon, B., and Barnett, T. A. (2018). The obese-asthma phenotype in children: An exacerbating situation? *J Allergy Clin Immunol*, 141(4):1239–1249.e4.

Luskin, A. T., Chipps, B. E., Rasouliyan, L., Miller, D. P., Haselkorn, T., and Dorenbaum, A. (2014). Impact of asthma exacerbations and asthma triggers on asthma-related quality of life in patients with severe or difficult-to-treat asthma. *J Allergy Clin Immunol Pract*, 2(5):544–552.e2.

Lutter, R. and Ravanetti, L. (2019). Cadherin-related family member 3 (CDHR3) drives differentiation of ciliated bronchial epithelial cells and facilitates rhinovirus C infection, although with a little help. *J Allergy Clin Immunol*, 144(4):926–927.

Lutz, S. M., Cho, M. H., Young, K., Hersh, C. P., Castaldi, P. J., McDonald, M.-L., Regan, E., Mattheisen, M., DeMeo, D. L., Parker, M., and et al. (2015). A genome-wide association study identifies risk loci for spirometric measures among smokers of European and African ancestry. *BMC Genet*, 16:138.

Maddala, R., Skiba, N. P., and Rao, P. V. (2017). Vertebrate lonesome kinase regulated extracellular matrix protein phosphorylation, cell shape, and adhesion in trabecular meshwork cells. *J Cell Physiol*, 232(9):2447–2460.

Major, S., Vézina, K., and Tse, S. M. (2021). Lung function of children following an intensive care unit admission for asthma. *Pediatr Allergy Immunol Pulmonol*, 34(1):1–6.

Mak, A. C., Kachuri, L., Hu, D., Eng, C., Huntsman, S., Elhawary, J., Gupta, N., Gabriel, S., Xiao, S., Gui, H., and et al. (2021). Gene expression in African Americans and Latinos reveals ancestry-specific patterns of genetic architecture. *BioRxiv*.

Malinovschi, A., Fonseca, J. A., Jacinto, T., Alving, K., and Janson, C. (2013). Exhaled nitric oxide levels and blood eosinophil counts independently associate with wheeze and asthma events in national health and nutrition examination survey subjects. *J Allergy Clin Immunol*, 132(4):821–827.e1–5.

Mandal, A. K., Zhang, Z., Ray, R., Choi, M. S., Chowdhury, B., Pattabiraman, N., and Mukherjee, A. B. (2004). Uteroglobin represses allergen-induced inflammatory response by blocking PGD2 receptor–mediated functions. *J Exp Med*, 199(10):1317–1330.

Maples, B. K., Gravel, S., Kenny, E. E., and Bustamante, C. D. (2013). RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. *Am J Hum Genet*, 93(2):278–288.

Martin, M. J., Beasley, R., and Harrison, T. W. (2020). Towards a personalised treatment approach for asthma attacks. *Thorax*, 75(12):1119–1129.

Masoli, M., Fabian, D., Holt, S., and Beasley, R. (2004). The global burden of asthma: executive summary of the GINA Dissemination Committee Report. *Allergy*, 59(5):469–478.

Mathias, R. A., Taub, M. A., Gignoux, C. R., Fu, W., Musharoff, S., O'Connor, T. D., Vergara, C., Torgerson, D. G., Pino-Yanes, M., Shringarpure, S. S., and et al. (2016). A continuum of admixture in the Western Hemisphere revealed by the African Diaspora genome. *Nat Commun*, 7:12522.

Matsunuma, R., Chan, D. W., Kim, B.-J., Singh, P., Han, A., Saltzman, A. B., Cheng, C., Lei, J. T., Wang, J., Roberto da Silva, L., and et al. (2018). DPYSL3 modulates mitosis, migration, and epithelial-to-mesenchymal transition in claudin-low breast cancer. *Proc Natl Acad Sci U S A*, 115(51):E11978–E11987.

Mayes, M. D., Bossini-Castillo, L., Gorlova, O., Martin, J. E., Zhou, X., Chen, W. V., Assassi, S., Ying, J., Tan, F. K., Arnett, F. C., and et al. (2014). Immunochip analysis identifies multiple susceptibility loci for systemic sclerosis. *Am J Hum Genet*, 94(1):47–61.

McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., Kang, H. M., Fuchsberger, C., Danecek, P., Sharp, K., and et al. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*, 48(10):1279–1283.

McCauley, K., Durack, J., Valladares, R., Fadrosh, D. W., Lin, D. L., Calatroni, A., LeBeau, P. K., Tran, H. T., Fujimura, K. E., LaMere, B., and et al. (2019). Distinct nasal airway bacterial microbiotas differentially relate to exacerbation in pediatric patients with asthma. *J Allergy Clin Immunol*, 144(5):1187–1197.

McGeachie, M. J., Wu, A. C., Tse, S. M., Clemmer, G. L., Sordillo, J., Himes, B. E., Lasky-Su, J., Chase, R. P., Martinez, F. D., Weeke, P., and et al. (2015). CTNNA3 and SEMA3D: promising loci for asthma exacerbation identified through multiple genome-wide association studies. *J Allergy Clin Immunol*, 136(6):1503–1510.

McKinstry, B., Sullivan, F. M., Vasishta, S., Armstrong, R., Hanley, J., Haughney, J., Philip, S., Smith, B. H., Wood, A., and Palmer, C. N. A. (2017). Cohort profile: the Scottish Research register SHARE. A register of people interested in research participation linked to NHS data sets. *BMJ Open*, 7(2):e013351.

McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*, 28:495–501.

Melén, E., Bergström, A., Kull, I., Almqvist, C., Andersson, N., Asarnoj, A., Borres, M. P., Georgellis, A., Pershagen, G., Westman, M., and et al. (2020). Male sex is strongly associated with IgE-sensitization to airborne but not food allergens: results up to age 24 years from the BAMSE birth cohort. *Clin Transl Allergy*, 10:15.

Mersha, T. B., Qin, K., Beck, A. F., Ding, L., Huang, B., and Kahn, R. S. (2021). Genetic ancestry differences in pediatric asthma readmission are mediated by socioenvironmental factors. *J Allergy Clin Immunol*, 148:1210–1218.e4.

Meurs, H., G.çosens, R., and Zaagsma, J. (2008). Airway hyperresponsiveness in asthma: lessons from in vitro model systems and animal models. *Eur Respir J*, 32(2):487–502.

Michigan Genomics Initiative (2018). MGI freeze2 PheWeb. Available from: http://pheweb.sph.umich.edu/pheno/495.2.

Mills, M. C. and Rahal, C. (2020). The GWAS Diversity Monitor tracks diversity by disease in real time. *Nat Genet*, 52(3):242–243.

Min, J. L., Hemani, G., Hannon, E., Dekkers, K. F., Castillo-Fernandez, J., Luijk, R., Carnero-Montoro, E., Lawson, D. J., Burrows, K., Suderman, M., and et al. (2021). Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat Genet*, 53(9):1311–1321.

Moore, J. E., Purcaro, M. J., Pratt, H. E., Epstein, C. B., Shoresh, N., Adrian, J., Kawli, T., Davis, C. A., Dobin, A., Kaul, R., and et al. (2020). Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*, 583(7818):699–710.

Mootz, M., Jakwerth, C. A., Schmidt-Weber, C. B., and Zissler, U. M. (2021). Secretoglobins in the big picture of immunoregulation in airway diseases. *Allergy*, Online ahead of print.

Murray, C. S., Jackson, D. J., and Teague, W. G. (2021). Prevention and outpatient treatment of asthma exacerbations in children. *J Allergy Clin Immunol Pract*, 9(7):2567–2576.

Muñoz, X., Álvarez Puebla, M. J., Arismendi, E., Arochena, L., Ausín, M. D. P., Barranco, P., Bobolea, I., Cañas, J. A., Cardaba, B., Crespo, A., and et al. (2018). The MEGA project: A study of the mechanisms involved in the genesis and disease course of asthma. Asthma cohort creation and long-term follow-up. *Arch Bronconeumol*, S0300-2896(18):30009–7.

Márquez, A., Kerick, M., Zhernakova, A., Gutierrez-Achury, J., Chen, W.-M., Onengut-Gumuscu, S., González-Álvaro, I., Rodriguez-Rodriguez, L., Rios-Fernández, R., González-Gay, M. A., and et al. (2018). Meta-analysis of Immunochip data of four autoimmune diseases reveals novel single-disease and cross-phenotype associations. *Genome Med*, 10(1):97.

Nadanaka, S., Hashiguchi, T., and Kitagawa, H. (2020). Aberrant glycosaminoglycan biosynthesis by tumor suppressor EXTL2 deficiency promotes liver inflammation and tumorigenesis through Toll-like 4 receptor signaling. *FASEB J*, 34(6):8385–8401.

Nadanaka, S., Zhou, S., Kagiyama, S., Shoji, N., Sugahara, K., Sugihara, K., Asano, M., and Kitagawa, H. (2013). EXTL2, a member of the EXT family of tumor suppressors, controls glycosaminoglycan biosynthesis in a xylose kinase-dependent manner. *J Biol Chem*, 288(13):9321–9333.

Nanishi, M., Press, V. G., Miller, J. B., Eastin, C., Aurora, T., Crocker, E., Fujiogi, M., Camargo, C. A., and Hasegawa, K. (2021). Hospital-initiated care bundle, posthospitalization care, and outcomes in adults with asthma exacerbation. *J Allergy Clin Immunol Pract*.

Navanandan, N., Hatoun, J., Celedón, J. C., and Liu, A. H. (2021). Predicting severe asthma exacerbations in children: Blueprint for today and tomorrow. *J Allergy Clin Immunol Pract*, 9(7):2619–2626.

Naveed, S. U. N., Clements, D., Jackson, D. J., Philp, C., Billington, C. K., Soomro, I., Reynolds, C., Harrison, T. W., Johnston, S. L., Shaw, D. E., and et al. (2017). Matrix metalloproteinase-1 activation contributes to airway smooth muscle growth and asthma severity. *Am J Respir Crit Care Med*, 195(8):1000–1009.

Neale lab (2018). GWAS round 2 results. Available from: http://www.nealelab.is/uk-biobank.

Neophytou, A. M., Oh, S. S., White, M. J., Mak, A. C. Y., Hu, D., Huntsman, S., Eng, C., Serebrisky, D., Borrell, L. N., Farber, H. J., and et al. (2018). Secondhand smoke exposure and asthma outcomes among African-American and Latino children with asthma. *Thorax*, 73(11):1041–1048.

Nieuwenhuis, M. A., Siedlinski, M., van den Berge, M., Granell, R., Li, X., Niens, M., van der Vlies, P., Altmüller, J., Nürnberg, P., Kerkhof, M., and et al. (2016). Combining genomewide association study and lung eQTL analysis provides evidence for novel genes associated with asthma. *Allergy*, 71(12):1712–1720.

Niimi, A., Matsumoto, H., Amitani, R., Nakano, Y., Mishima, M., Minakuchi, M., Nishimura, K., Itoh, H., and Izumi, T. (2000). Airway wall thickness in asthma assessed by Computed Tomography. *Am J Respir Crit Care Med*, 162(4):1518–1523.

Nishimura, K. K., Galanter, J. M., Roth, L. A., Oh, S. S., Thakur, N., Nguyen, E. A., Thyne, S., Farber, H. J., Serebrisky, D., Kumar, R., and et al. (2013). Early-life air pollution and asthma risk in minority children. The GALA II and SAGE II studies. *Am J Respir Crit Care Med*, 188(3):309–318.

Nunes, C., Pereira, A. M., and Morais-Almeida, M. (2017). Asthma costs and social impact. *Asthma Res Pract*, 3(1):1.

Ober, C. and Hoffjan, S. (2006). Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun*, 7(2):95–100.

Oh, C., Verma, A., and Aachoui, Y. (2020). Caspase-11 non-canonical inflammasomes in the lung. *Front Immunol*, 11:1895.

Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki, A., Yoshida, S., and et al. (2014). Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature*, 506(7488):376–381.

Oliveira, P., Costa, G. N. O., Damasceno, A. K. A., Hartwig, F. P., Barbosa, G. C. G., Figueiredo, C. A., Ribeiro-Silva, R. d. C., Pereira, A., Lima-Costa, M. F., Kehdy, F. S., and et al. (2018). Genome-wide burden and association analyses implicate copy number variations in asthma risk among children and young adults from latin america. *Sci Rep*, 8(1):14475.

Ongen, H., Buil, A., Brown, A. A., Dermitzakis, E. T., and Delaneau, O. (2016). Fast and efficient QTL mapper for thousands of molecular phenotypes. *Bioinformatics*, 32(10):1479–85.

Oraka, E., Iqbal, S., Flanders, W. D., Brinker, K., and Garbe, P. (2013). Racial and ethnic disparities in current asthma and emergency department visits: findings from the National Health Interview Survey, 2001-2010. *J Asthma*, 50(5):488–496.

Oren, E. and Martinez, F. D. (2020). Stress and asthma. *Ann Allergy Asthma Immunol*, 125(4):372–373.e1.

Ortega, V. E., Hawkins, G. A., Moore, W. C., Hastie, A. T., Ampleford, E. J., Busse, W. W., Castro, M., Chardon, D., Erzurum, S. C., Israel, E., and et al. (2014). Effect of rare variants in ADRB2 on risk of severe exacerbations and symptom control during long acting β-agonist treatment in a multiethnic asthma population: a genetic study. *Lancet Respir Med*, 2(3):204–213.

Oscanoa, J., Sivapalan, L., Gadaleta, E., Ullah, A. Z. D., Lemoine, N. R., and Chelala, C. (2020). SNPnexus: a web server for functional annotation of human genome sequence variation (2020 update). *Nucleic Acids Res*, 48(W1):W185–W192.

O'Brian, A. L., Lemanske, R. F., Evans, M. D., Gangnon, R. E., Gern, J. E., and Jackson, D. J. (2012). Recurrent severe exacerbations in early life and reduced lung function at school age. *J Allergy Clin Immunol*, 129(4):1162–1164.

O'Byrne, P. M., Pedersen, S., Lamm, C. J., Tan, W. C., Busse, W. W., and Group, S. I. (2009). Severe exacerbations and decline in lung function in asthma. *Am J Respir Crit Care Med*, 179(1):19–24.

Pai, A. A., Pritchard, J. K., and Gilad, Y. (2015). The genetic and mechanistic basis for variation in gene regulation. *PLoS Genet*, 11(1):e1004857.

Palmer, C. N. A., Doney, A. S. F., Ismail, T., Lee, S. P., Murrie, I., Macgregor, D. F., and Mukhopadhyay, S. (2007). PPARG locus haplotype variation and exacerbations in asthma. *J Clin Pharm Ther*, 81(5):713–718.

Palmer, C. N. A., Doney, A. S. F., Lee, S. P., Murrie, I., Ismail, T., Macgregor, D. F., and Mukhopadhyay, S. (2006a). Glutathione S-transferase M1 and P1 genotype, passive smoking, and peak expiratory flow in asthma. *Pediatrics*, 118(2):710–716.

Palmer, C. N. A., Lipworth, B. J., Lee, S., Ismail, T., Macgregor, D. F., and Mukhopadhyay, S. (2006b). Arginine-16 beta2 adrenoceptor genotype predisposes to exacerbations in young asthmatics taking regular salmeterol. *Thorax*, 61(11):940–944.

Pandey, R. C., Michel, S., Schieck, M., Binia, A., Liang, L., Klopp, N., Franke, A., von Berg, A., Bufe, A., Rietschel, E., and et al. (2013). Polymorphisms in extracellular signal-regulated kinase family influence genetic susceptibility to asthma. *J Allergy Clin Immunol*, 131(4):1245–1247.

Papadopoulos, N. G., Christodoulou, I., Rohde, G., Agache, I., Almqvist, C., Bruno, A., Bonini, S., Bont, L., Bossios, A., Bousquet, J., and et al. (2011). Viruses and bacteria in acute asthma exacerbations-a GA^2 LEN-DARE systematic review. *Allergy*, 66(4):458–468.

Parker, M. M., Foreman, M. G., Abel, H. J., Mathias, R. A., Hetmanski, J. B., Crapo, J. D., Silverman, E. K., Beaty, T. H., and COPDGene Investigators (2014). Admixture mapping identifies a quantitative trait locus associated with FEV1/FVC in the COPDGene Study. *Genet Epidemiol*, 38(7):652–659.

Patel, M., Pilcher, J., Reddel, H. K., Qi, V., Mackey, B., Tranquilino, T., Shaw, D., Black, P., Weatherall, M., and Beasley, R. (2014). Predictors of severe exacerbations, poor asthma control, and β-agonist overuse for patients with asthma. *J Allergy Clin Immunol Pract*, 2(6):751–758.e1.

Perez-Garcia, J., Hernández-Pérez, J. M., González-Pérez, R., Sardón, O., Martin-Gonzalez, E., Espuela-Ortiz, A., Mederos-Luis, E., Callero, A., Herrera-Luis, E., Corcuera, P., and et al. (2020). The Genomics and Metagenomics of Asthma Severity (GEMAS) study: Rationale and design. *J Pers Med*, 10:123.

Pino-Yanes, M., Gignoux, C. R., Galanter, J. M., Levin, A. M., Campbell, C. D., Eng, C., Huntsman, S., Nishimura, K. K., Gourraud, P.-A., Mohajeri, K., and et al. (2015a). Genome-wide association study and admixture mapping reveal new loci associated with total IgE levels in Latinos. *J Allergy Clin Immunol*, 135(6):1502–1510.

Pino-Yanes, M., Thakur, N., Gignoux, C. R., Galanter, J. M., Roth, L. A., Eng, C., Nishimura, K. K., Oh, S. S., Vora, H., Huntsman, S., and et al. (2015b). Genetic ancestry influences asthma susceptibility and lung function among Latinos. *J Allergy Clin Immunol*, 135(1):228–235.

Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, 6(1):7–11.

Price, D. B., Rigazio, A., Campbell, J. D., Bleecker, E., Corrigan, C. J., Thomas, M., Wenzel, S. E., Wilson, A. M., Small, M. B., Gopalan, G., and et al. (2015). Blood eosinophil count and prospective annual asthma disease burden: a UK cohort study. *Lancet Respir Med*, 3(11):849–858.

Pu, A., Mishra, M. K., Dong, Y., Ghorbanigazar, S., Stephenson, E. L., Rawji, K. S., Silva, C., Kitagawa, H., Sawcer, S., and Yong, V. W. (2020). The glycosyltransferase EXTL2 promotes proteoglycan deposition and injurious neuroinflammation following demyelination. *J Neuroinflammation*, 17(1):220.

Puranik, S., Forno, E., Bush, A., and Celedón, J. C. (2017). Predicting severe asthma exacerbations in children. *Am J Respir Crit Care Med*, 195(7):854–859.

Quanjer, P. H., Stanojevic, S., Cole, T. J., Baur, X., Hall, G. L., Culver, B. H., Enright, P. L., Hankinson, J. L., Ip, M. S., Zheng, J., and et al. (2012). Multi-ethnic reference values for spirometry for the 3-95-yr age range: The global lung function 2012 equations. *Eur Respir J*, 40(6):1324–1343.

Quick, C., Anugu, P., Musani, S., Weiss, S. T., Burchard, E. G., White, M. J., Keys, K. L., Cucca, F., Sidore, C., Boehnke, M., and et al. (2020). Sequencing and imputation in GWAS: Cost-effective strategies to increase power and genomic coverage across diverse populations. *Genet Epidemiol*, 44(6):537–549.

Racine, G., Forget, A., Moullec, G., Jiao, T., Blais, L., and Lemiere, C. (2021). Predictors of asthma control and exacerbations: A real-world study. *J Allergy Clin Immunol Pract*, 9(7):2802–2811.e2.

Rahmani, E., Yedidim, R., Shenhav, L., Schweiger, R., Weissbrod, O., Zaitlen, N., and Halperin, E. (2017). GLINT: A user-friendly toolset for the analysis of high-throughput DNA-methylation array data. *Bioinformatics*, 33(12):1870–1872.

Rahmani, E., Zaitlen, N., Baran, Y., Eng, C., Hu, D., Galanter, J., Oh, S., Burchard, E. G., Eskin, E., Zou, J., and et al. (2016). Sparse PCA corrects for cell type heterogeneity in epigenome-wide association studies. *Nat Methods*, 13(5):443–5.

Ramadan, A. A., Gaffin, J. M., Israel, E., and Phipatanakul, W. (2019). Asthma and corticosteroid responses in childhood and adult asthma. *Clin Chest Med*, 40(1):163–177.

Ramaswamy, G., Karim, M. A., Murti, K. G., and Jackowski, S. (2004). PPARα controls the intracellular coenzyme A concentration via regulation of PANK1α gene expresion. *J Lipid Res*, 45(1):17–31.

Ramsahai, J. M., Hansbro, P. M., and Wark, P. A. B. (2019). Mechanisms and management of asthma exacerbations. *Am J Respir Crit Care Med*, 199(4):423–432.

Rastogi, D., Madhok, N., and Kipperman, S. (2013). Caregiver asthma knowledge, aptitude, and practice in high healthcare utilizing children: Effect of an educational intervention. *Pediatr Allergy Immunol Pulmonol*, 26(3):128–139.

Reddel, H. K., Taylor, D. R., Bateman, E. D., Boulet, L. P., Boushey, H. a., Busse, W. W., Casale, T. B., Chanez, P., Enright, P. L., Gibson, P. G., and et al. (2009). An official American Thoracic Society/European Respiratory Society statement: asthma control and exacerbations: standardizing endpoints for clinical asthma trials and clinical practice. *Am J Respir Crit Care Med*, 180(1):59–99.

Revollo, L. D., Merrill-Skoloff, G., De Ceunynck, K., Dilks, J. R., Guo, S., Bordoli, M. R., Peters, C. G., Noetzli, L., Ionescu, A. M., Rosen, V., and et al. (2021). The secreted tyrosine kinase VLK is essential for normal platelet activation and thrombus formation. *Blood*, 139(1):104–117.

Ritz, T., Wittchen, H.-U., Klotsche, J., Mühlig, S., Riedel, O., and sap-NEEDs study group * (2016). Asthma trigger reports are associated with low quality of life, exacerbations, and emergency treatments. *Ann Am Thorac Soc*, 13(2):204–211.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1):139–140.

Rosser, F. J., Forno, E., Cooper, P. J., and Celedón, J. C. (2014). Asthma in Hispanics: An 8-year update. *Am J Respir Crit Care Med*, 189(11):1316–1327.

Rothenberg, M. E. and Hogan, S. P. (2006). The eosinophil. *Annu Rev Immunol*, 24(1):147–174.

Royce, S. G., Cheng, V., Samuel, C. S., and Tang, M. L. (2012). The regulation of fibrosis in airway remodeling in asthma. *Mol Cell Endocrinol*, 351(2):167–175.

Rumpel, J. A., Ahmedani, B. K., Peterson, E. L., Wells, K. E., Yang, M., Levin, A. M., Yang, J. J., Kumar, R., Burchard, E. G., and Williams, L. K. (2012). Genetic ancestry and its association with asthma exacerbations among African American subjects with asthma. *J Allergy Clin Immunol*, 130(6):1302–1306.

Sakaue, S., Kanai, M., Tanigawa, Y., Karjalainen, J., Kurki, M., Koshiba, S., Narita, A., Konuma, T., Yamamoto, K., Akiyama, M., and et al. (2021). A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet*, 53(10):1415–1424.

Salciccioli, J. D., She, L., Tulchinsky, A., Rockhold, F., Cardet, J. C., and Israel, E. (2021). Effect of COVID-19 on asthma exacerbation. *J Allergy Clin Immunol Pract*, 9(7):2896–2899.e1.

Sayers, E. W., Beck, J., Bolton, E. E., Bourexis, D., Brister, J. R., Canese, K., Comeau, D. C., Funk, K., Kim, S., Klimke, W., and et al. (2021). Database resources of the national center for biotechnology information. *Nucleic Acids Res*, 49(D1):D10–D17.

Schatz, M., Meckley, L. M., Kim, M., Stockwell, B. T., and Castro, M. (2014). Asthma exacerbation rates in adults are unchanged over a 5-year period despite high-intensity therapy. *J Allergy Clin Immunol Pract*, 2(5):570–574.e1.

Schatz, M., Zeiger, R. S., Zhang, F., Chen, W., Yang, S.-J., and Camargo, C. A. (2013). Overweight/obesity and risk of seasonal asthma exacerbations. *J Allergy Clin Immunol Pract*, 1(6):618–622.

Schieck, M., Michel, S., Suttner, K., Illig, T., Zeilinger, S., Franke, A., Vogelberg, C., Von Berg, A., Bufe, A., Heinzmann, A., and et al. (2014). Genetic variation in TH17 pathway genes, childhood asthma, and total serum IgE levels. *J Allergy Clin Immunol*, 133(3):888–891.

Schoettler, N., Rodríguez, E., Weidinger, S., and Ober, C. (2019). Advances in asthma and allergic disease genetics: Is bigger always better? *J Allergy Clin Immunol*, 144(6):1495–1506.

Schofield, E. C., Carver, T., Achuthan, P., Freire-Pritchett, P., Spivakov, M., Todd, J. A., and Burren, O. S. (2016). CHiCP: a web-based tool for the integrative and interactive visualization of promoter capture Hi-C datasets. *Bioinformatics*, 32(16):2511–2513.

Sears, M. R. (2008). Epidemiology of asthma exacerbations. *J Allergy Clin Immunol*, 122(4):662–668.

Sears, M. R. (2014). Trends in the prevalence of asthma. *Chest*, 145(2):219–225.

Semprini, R., Williams, M., Semprini, A., McDouall, A., Fingleton, J., Holweg, C., Weatherall, M., Beasley, R., and Braithwaite, I. (2018). Type 2 biomarkers and prediction of future exacerbations and lung function decline in adult asthma. *J Allergy Clin Immunol Pract*, 6(6):1982–1988.e1.

Shabalin, A. A. (2012). Matrix eQTL: Ultra fast eQTL analysis via large matrix operations. *Bioinformatics*, 28(10):1353–1358.

Shah, R. and Newcomb, D. C. (2018). Sex bias in asthma prevalence and pathogenesis. *Front Immunol*, 9:2997.

Shah, S. A., Quint, J. K., Nwaru, B. I., and Sheikh, A. (2021). Impact of COVID-19 national lockdown on asthma exacerbations: interrupted time-series analysis of English primary care data. *Thorax*, 76(9):860–866.

Sharma, S., Raby, B. A., Hunninghake, G. M., Soto-Quirós, M., Avila, L., Murphy, A. J., Lasky-Su, J., Klanderman, B. J., Sylvia, J. S., Weiss, S. T., and et al. (2009). Variants in TGFB1, dust mite exposure, and disease severity in children with asthma. *Am J Respir Crit Care Med*, 179(5):356–362.

Shaw, D. E., Sousa, A. R., Fowler, S. J., Fleming, L. J., Roberts, G., Corfield, J., Pandis, I., Bansal, A. T., Bel, E. H., Auffray, C., and et al. (2015). Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma cohort. *Eur Respir J*, 46(5):1308–1321.

Shi, D., Sheng, A., and Chi, L. (2021). Glycosaminoglycan-protein interactions and their roles in human disease. *Front Mol Biosci*, 8:639666.

Shi, G., Abbott, K. N., Wu, W., Salter, R. D., and Keyel, P. A. (2017). Dnase1L3 regulates inflammasome-dependent cytokine secretion. *Front Immunol*, 8:522.

Shrimanker, R., Pavord, I. D., Yancey, S., Heaney, L. G., Green, R. H., Bradding, P., Hargadon, B., Brightling, C. E., Wardlaw, A. J., and Haldar, P. (2018). Exacerbations of severe asthma in patients treated with mepolizumab. *Eur Respir J*, 52(6):1801127.

Shrine, N., Portelli, M. A., John, C., Soler Artigas, M., Bennett, N., Hall, R., Lewis, J., Henry, A. P., Billington, C. K., Ahmad, A., and et al. (2019). Moderate-to-severe asthma in individuals of European ancestry: a genome-wide association study. *Lancet Respir Med*, 7(1):20–34.

Shriner, D. (2017). Overview of admixture mapping. *Curr Protoc Hum Genet*, 94:1.23.1–1.23.8.

Shumyatcher, M., Hong, R., Levin, J., and Himes, B. E. (2017). Disease-specific integration of omics data to guide functional validation of genetic associations. *AMIA Annu Symp Proc*, 2017:1589–1596.

Silber, J. H., Rosenbaum, P. R., Calhoun, S. R., Reiter, J. G., Hill, A. S., Guevara, J. P., Zorc, J. J., and Even-Shoshan, O. (2017). Racial disparities in medicaid asthma hospitalizations. *Pediatrics*, 139(1):e20161221.

Silver, J. D., Sutherland, M. F., Johnston, F. H., Lampugnani, E. R., McCarthy, M. A., Jacobs, S. J., Pezza, A. B., and Newbigin, E. J. (2018). Seasonal asthma in Melbourne, Australia, and some observations on the occurrence of thunderstorm asthma and its predictability. *PloS One*, 13(4):e0194929.

Sio, Y. Y., Anantharaman, R., Lee, S. Q. E., Matta, S. A., Ng, Y. T., and Chew, F. T. (2020a). The asthma-associated PER1-like domain-containing protein 1 (PERLD1) haplotype influences Soluble Glycosylphosphatidylinositol Anchor Protein (sGPI-AP) levels in serum and immune cell proliferation. *Sci Rep*, 10(1):715.

Sio, Y. Y., Matta, S. A., Ng, Y. T., and Chew, F. T. (2020b). Epistasis between phenylethanolamine n-methyltransferase and β2-adrenergic receptor influences extracellular epinephrine level and associates with the susceptibility to allergic asthma. *Clin Exp Allergy*, 50(3):352–363.

Sisirak, V., Sally, B., D'Agati, V., Martinez-Ortiz, W., Özçakar, Z. B., David, J., Rashidfarrokhi, A., Yeste, A., Panea, C., Chida, A. S., and et al. (2016). Digestion of chromatin in apoptotic cell microparticles prevents autoimmunity. *Cell*, 166(1):88–101.

Slob, E. M. A., Richards, L. B., Vijverberg, S. J. H., Longo, C., Koppelman, G. H., Pijnenburg, M. W. H., Bel, E. H. D., Neerincx, A. H., Herrera Luis, E., Perez-Garcia, J., and et al. (2021). Genome-wide association studies of exacerbations in children using long-acting beta2-agonists. *Pediatr Allergy Immunol*, 32(6):1197–1207.

Soler Artigas, M., Wain, L. V., Miller, S., Kheirallah, A. K., Huffman, J. E., Ntalla, I., Shrine, N., Obeidat, M., Trochet, H., McArdle, W. L., and et al. (2015). Sixteen new lung function signals identified through 1000 Genomes Project reference panel imputation. *Nat Commun*, 6:8658.

Song, H. J., Blake, K. V., Wilson, D. L., Winterstein, A. G., and Park, H. (2020). Medical costs and productivity loss due to mild, moderate, and severe asthma in the United States. *J Asthma Allergy*, 13:545–555.

Song, H. J., Blake, K. V., Wilson, D. L., Winterstein, A. G., and Park, H. (2021). Health-related quality of life and health utilities of mild, moderate, and severe asthma: Evidence from the Medical Expenditure Panel Survey. *J Asthma Allergy*, 14:929–941.

Sordillo, J. E., Kelly, R., Bunyavanich, S., McGeachie, M., Qiu, W., Croteau-Chonka, D. C., Soto-Quiros, M., Avila, L., Celedón, J. C., Brehm, J. M., and et al. (2015). Genome-wide expression profiles identify potential targets for gene-environment interactions in asthma severity. *J Allergy Clin Immunol*, 136(4):885–892.e2.

Staley, J. R., Blackshaw, J., Kamat, M. A., Ellis, S., Surendran, P., Sun, B. B., Paul, D. S., Freitag, D., Burgess, S., Danesh, J., and et al. (2016). PhenoScanner: a database of human genotype–phenotype associations. *Bioinformatics*, 32(20):3207–3209.

Stegle, O., Parts, L., Durbin, R., and Winn, J. (2010). A Bayesian framework to account for complex non-genetic factors in gene expression levels greatly increases power in eQTL studies. *PLoS Comput Biol*, 6(5):e1000770.

Strid, J. M. C., Gammelager, H., Johansen, M. B., Tønnesen, E., and Christiansen, C. F. (2013). Hospitalization rate and 30-day mortality among patients with status asthmaticus in Denmark: a 16-year nationwide population-based cohort study. *Clin Epidemiol*, 5:345–355.

Su, Y.-T., Gao, C., Liu, Y., Guo, S., Wang, A., Wang, B., Erdjument-Bromage, H., Miyagi, M., Tempst, P., and Kao, H.-Y. (2013). Monoubiquitination of filamin B regulates vascular endothelial growth factor-mediated trafficking of histone deacetylase 7. *Mol Cell Biol*, 33(8):1546–1560.

Suarez-Pajes, E., Díaz-de Usera, A., Marcelino-Rodríguez, I., Guillen-Guio, B., and Flores, C. (2021). Genetic ancestry inference and its application for the genetic mapping of human diseases. *Int J Mol Sci*, 22(13):6962.

Sun, J., Huang, N., Ma, W., Zhou, H., and Lai, K. (2019). Protective effects of metformin on lipopolysaccharide-induced airway epithelial cell injury via NF-κB signaling inhibition. *Mol Med Rep*, 19(3):1817–1823.

Sunadome, H., Matsumoto, H., Petrova, G., Kanemitsu, Y., Tohda, Y., Horiguchi, T., Kita, H., Kuwabara, K., Tomii, K., Otsuka, K., and et al. (2017). IL4R and ADAM33 as genetic markers in asthma exacerbations and type-2 inflammatory endotype. *Clin Exp Allergy*, 47(8):998–1006.

Suruki, R. Y., Daugherty, J. B., Boudiaf, N., and Albers, F. C. (2017). The frequency of asthma exacerbations and healthcare utilization in patients with asthma from the UK and USA. *BMC Pulm Med*, 17(1):74.

Syabbalo, N. (2020). Neutrophilic asthma: a complex phenotype of severe asthma. *J Lung Pulm Respir Res*, 7(1):18–24.

Sze, E., Bhalla, A., and Nair, P. (2020). Mechanisms and therapeutic strategies for non-T2 asthma. *Allergy*, 75(2):311–325.

Taliun, D., Harris, D. N., Kessler, M. D., Carlson, J., Szpiech, Z. A., Torres, R., Taliun, S. A. G., Corvelo, A., Gogarten, S. M., Kang, H. M., and et al. (2021). Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature*, 590(7845):290–299.

Tantisira, K. G., Silverman, E. S., Mariani, T. J., Xu, J., Richter, B. G., Klanderman, B. J., Litonjua, A. A., Lazarus, R., Rosenwasser, L. J., Fuhlbrigge, A. L., and et al. (2007). FCER2: A pharmacogenetic basis for severe exacerbations in children with asthma. *J Allergy Clin Immunol*, 120(6):1285–1291.

Tattersfield, A. E., Postma, D. S., Barnes, P. J., Svensson, K., Bauer, C. A., O'Byrne, P. M., Löfdahl, C. G., Pauwels, R. A., and Ullman, A. (1999). Exacerbations of asthma: a descriptive study of 425 severe exacerbations. The FACET International Study Group. *Am J Respir Crit Care Med*, 160(2):594–549.

Tavendale, R., Macgregor, D. F., Mukhopadhyay, S., and Palmer, C. N. (2008). A polymorphism controlling ORMDL3 expression is associated with asthma that is poorly controlled by current medications. *J Allergy Clin Immunol*, 121(4):860–863.

Taylor, D. R., Drazen, J. M., Herbison, G. P., Yandava, C. N., Hancox, R. J., and Town, G. I. (2000). Asthma exacerbations during long term beta agonist use: influence of beta(2) adrenoceptor polymorphism. *Thorax*, 55(9):762–767.

Team, R. C. (2020). R: A language and environment for statistical computing.

Terao, C., Kawaguchi, T., Dieude, P., Varga, J., Kuwana, M., Hudson, M., Kawaguchi, Y., Matucci-Cerinic, M., Ohmura, K., Riemekasten, G., and et al. (2017). Transethnic meta-analysis identifies GSDMA and PRDM1 as susceptibility genes to systemic sclerosis. *Ann Rheum Dis*, 76(6):1150–1158.

Thakur, N., Oh, S. S., Nguyen, E. A., Martin, M., Roth, L. A., Galanter, J., Gignoux, C. R., Eng, C., Davis, A., Meade, K., and et al. (2013). Socioeconomic status and childhood asthma in urban minority youths. The GALA II and SAGE II studies. *Am J Respir Crit Care Med*, 188(10):1202–1209.

To, T., Stanojevic, S., Moores, G., Gershon, A. S., Bateman, E. D., Cruz, A. A., and Boulet, L.-P. (2012). Global asthma prevalence in adults: findings from the cross-sectional world health survey. *BMC Public Health*, 12:204.

Trawick, D. R., Holm, C., and Wirth, J. (2001). Influence of gender on rates of hospitalization, hospital course, and hypercapnea in high-risk patients admitted for asthma: a 10-year retrospective study at Yale-New Haven Hospital. *Chest*, 119(1):115–119.

Trivedi, M. and Denton, E. (2019). Asthma in children and adults-what are the differences and what can they tell us about asthma? *Front Pediatr*, 7:256.

Tsai, C.-H., Wu, A. C., Chiang, B.-L., Yang, Y.-H., Hung, S.-P., Su, M.-W., Chang, Y.-J., and Lee, Y. L. (2020a). CEACAM3 decreases asthma exacerbations and modulates respiratory syncytial virus latent infection in children. *Thorax*, 75(9):725–734.

Tsai, Y.-M., Wu, K.-L., Chang, Y.-Y., Hung, J.-Y., Chang, W.-A., Chang, C.-Y., Jian, S.-F., Tsai, P.-H., Huang, Y.-C., Chong, I.-W., and et al. (2020b). Upregulation of Thr/Tyr kinase increases the cancer progression by neurotensin and Dihydropyrimidinase-Like 3 in lung cancer. *Int J Mol Sci*, 21(5):1640.

Tse, S. M., Krajinovic, M., Chauhan, B. F., Zemek, R., Gravel, J., Chalut, D., Poonai, N., Quach, C., Laberge, S., Ducharme, F. M., and research group of the Pediatric Emergency Research in Canada (PERC) network, D. (2019). Genetic determinants of acute asthma therapy response in children with moderate-to-severe asthma exacerbations. *Pediatr Pulmonol*, 54(4):378–385.

Turner, S., Francis, B., Vijverberg, S., Pino-Yanes, M., Maitland-van der Zee, A. H., Basu, K., Bignell, L., Mukhopadhyay, S., Tavendale, R., Palmer, C., and et al. (2016). Childhood asthma exacerbations and the arg16 β2-receptor polymorphism: A meta-analysis stratified by treatment. *J Allergy Clin Immunol*, 138(1):107–113.e5.

Turner, S., Francis, B., Wani, N., Vijverberg, S., Pino-Yanes, M., Mukhopadhyay, S., Tavendale, R., Palmer, C., Burchard, E. G., Merid, S. K., and et al. (2018). Variants in genes coding for glutathione S-transferases and asthma outcomes in children. *Pharmacogenomics*, 19(8):707–713.

Turner, S. W., Ayres, J. G., Macfarlane, T. V., Mehta, A., Mehta, G., Palmer, C. N., Cunningham, S., Adams, T., Aniruddhan, K., Bell, C., and et al. (2010). A methodology to establish a database to study gene environment interactions for childhood asthma. *BMC Med Res Methodol*, 10:107.

UK Biobank (2020). UKBiobank ICD PheWeb: Analysis of 1403 ICD-based traits using SAIGE. Available from: http://pheweb.sph.umich.edu/SAIGE-UKB/pheno/495.2.

U.S. National Heart, Lung, and Blood Institute of the National Institutes of Health (2007). *Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma*. Department of Health and Human Services, National Institutes of Health.

Utidjian, L. H., Fiks, A. G., Localio, A. R., Song, L., Ramos, M. J., Keren, R., Bell, L. M., and Grundmeier, R. W. (2017). Pediatric asthma hospitalizations among urban minority children and the continuity of primary care. *J Asthma*, 54(10):1051–1058.

Wang, X., Tanino, Y., Sato, S., Nikaido, T., Misa, K., Fukuhara, N., Fukuhara, A., Saito, J., Yokouchi, H., Ishida, T., and et al. (2015). Secretoglobin 3A2 attenuates lipopolysaccharide-induced inflammation through inhibition of ERK and JNK pathways in bronchial epithelial cells. *Inflammation*, 38(2):828–834.

Wang, Z., Ji, N., Chen, Z., Wu, C., Sun, Z., Yu, W., Hu, F., Huang, M., and Zhang, M. (2019). Next generation sequencing for long non-coding RNAs profile for CD4+ T cells in the mouse model of acute asthma. *Front Genet*, 10:545.

Ward, L. D. and Kellis, M. (2016). HaploReg v4: Systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res*, 44(D1):D877–D881.

Watanabe, K., Taskesen, E., Van Bochoven, A., and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*, 8(1):1–11.

Wenzel, S., Castro, M., Corren, J., Maspero, J., Wang, L., Zhang, B., Pirozzi, G., Sutherland, E. R., Evans, R. R., Joish, V. N., and et al. (2016). Dupilumab efficacy and safety in adults with uncontrolled persistent asthma despite use of medium-to-high-dose inhaled corticosteroids plus a long-acting $\beta 2$ agonist: a randomised double-blind placebo-controlled pivotal phase 2b dose-ranging trial. *Lancet*, 388(10039):31–44.

Winkler, T. and Frey, U. (2021). Airway remodeling: Shifting the trigger point for exacerbations in asthma. *J Allergy Clin Immunol*, 148(3):710–712.

Wittchen, E. S. (2009). Endothelial signaling in paracellular and transcellular leukocyte transmigration. *Front Biosci (Landmark Ed)*, 14:2522–2545.

Wu, A. C., Lasky-Su, J., Rogers, C. A., Klanderman, B. J., and Litonjua, A. A. (2010). Fungal exposure modulates the effect of polymorphisms of chitinases on emergency department visits and hospitalizations. *Am J Respir Crit Care Med*, 182(7):884–889.

Wu, A. C., Tantisira, K., Li, L., Schuemann, B., Weiss, S. T., Fuhlbrigge, A. L., and Group, C. A. M. P. R. (2011). Predictors of symptoms are different from predictors of severe exacerbations from asthma in children. *Chest*, 140(1):100–107.

Xu, Q., Wu, N., Cui, L., Wu, Z., and Qiu, G. (2017). Filamin B: The next hotspot in skeletal research? *J Genet Genomics*, 44(7):335–342.

Xu, Z., Niu, L., Li, L., and Taylor, J. A. (2016). ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. *Nucleic Acids Res*, 44(3):e20.

Yan, Q., Brehm, J., Pino-Yanes, M., Forno, E., Lin, J., Oh, S. S., Acosta-Perez, E., Laurie, C. C., Cloutier, M. M., Raby, B. A., and et al. (2017). A meta-analysis of genome-wide association studies of asthma in Puerto Ricans. *Eur Respir J*, 49(5):1601505.

Yan, Q., Forno, E., Herrera-Luis, E., Pino-Yanes, M., Qi, C., Rios, R., Han, Y.-Y., Kim, S., Oh, S., Acosta-Pérez, E., and et al. (2021a). A genome-wide association study of severe asthma exacerbations in Latino children and adolescents. *Eur Respir J*, 57(4):2002693.

Yan, Q., Forno, E., Herrera-Luis, E., Pino-Yanes, M., Yang, G., Oh, S., Acosta-Pérez, E., Hu, D., Eng, C., Huntsman, S., and et al. (2021b). A genome-wide association study of asthma hospitalizations in adults. *J Allergy Clin Immunol*, 147(3):933–940.

Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *Am J Hum Genet*, 88(1):76–82.

Ye, Y., Zhang, Z., Liu, Y., Diao, L., and Han, L. (2020). A multi-omics perspective of quantitative trait loci in precision medicine. *Trends Genet*, 36(5):318–336.

Yeung, T. Y., Ewing, C., Malanowska, A., Zuberbuhler, P., Balcom, M., Liu, J., and Amirav, I. (2018). Home management of childhood asthma exacerbations. *Pulm Ther*, 4(2):149–157.

Yokoyama, A. (2019). *Advances in Asthma: Pathophysiology, Diagnosis and Treatment*. Respiratory Disease Series: Diagnostic Tools and Disease Managements. Springer Singapore Pte. Limited, Singapore.

Yokoyama, S., Cai, Y., Murata, M., Tomita, T., Yoneda, M., Xu, L., Pilon, A. L., Cachau, R. E., and Kimura, S. (2018). A novel pathway of LPS uptake through syndecan-1 leading to pyroptotic cell death. *eLife*, 7.

Yokoyama, S., Nakayama, S., Xu, L., Pilon, A. L., and Kimura, S. (2021). Secretoglobin 3A2 eliminates human cancer cells through pyroptosis. *Cell Death Discov*, 7(1):12.

Zhan, X., Hu, Y., Li, B., Abecasis, G. R., and Liu, D. J. (2016). Rvtests: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics*, 32(9):1423–1426.

Zhao, Q., Yang, C., Wang, J., Li, Y., and Yang, P. (2017). Serum level of DNase1l3 in patients with dermatomyositis/polymyositis, systemic lupus erythematosus and rheumatoid arthritis, and its association with disease activity. *Clin Exp Med*, 17(4):459–465.

Zheng, J., Erzurumluoglu, A. M., Elsworth, B. L., Kemp, J. P., Howe, L., Haycock, P. C., Hemani, G., Tansey, K., Laurin, C., Early Genetics and Lifecourse Epidemiology (EAGLE) Eczema Consortium, and et al. (2017). LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics (Oxford, England)*, 33(2):272–279.

Zheng, Z., Huang, D., Wang, J., Zhao, K., Zhou, Y., Guo, Z., Zhai, S., Xu, H., Cui, H., Yao, H., and et al. (2020). QTLbase: an integrative resource for quantitative trait loci across multiple human molecular phenotypes. *Nucleic Acids Res*, 48(D1):D983–D991.

Zhou, Y., Jackson, D., Bacharier, L. B., Mauger, D., Boushey, H., Castro, M., Durack, J., Huang, Y., Lemanske, R. F., Storch, G. A., and et al. (2019). The upper-airway microbiota and loss of asthma control among asthmatic children. *Nat Commun*, 10(1):5714.

Ziyatdinov, A., Parker, M. M., Vaysse, A., Beaty, T. H., Kraft, P., Cho, M. H., and Aschard, H. (2020). Mixed-model admixture mapping identifies smoking-dependent loci of lung function in African Americans. *Eur J Hum Genet*, 28(5):656–668.

Zuurhout, M. J. L., Vijverberg, S. J. H., Raaijmakers, J. A. M., Koenderman, L., Postma, D. S., Koppelman, G. H., and Maitland-van der Zee, A. H. (2013). Arg16 ADRB2 genotype increases the risk of asthma exacerbation in children with a reported use of long-acting β2-agonists: results of the PACMAN cohort. *Pharmacogenomics*, 14:1965–1971.