8-9-2014

ROLE OF ALLERGIC SENSITIZATION, FILAGGRIN VARIANTS, AND DNA

Ali H. Ziyab
University of South Carolina - Columbia

ROLE OF ALLERGIC SENSITIZATION, FILAGGRIN VARIANTS, AND DNA METHYLATION ON THE RISK OF ALLERGIC DISORDERS

by

Ali H. Ziyab

Bachelor of Science
Washington State University, 2005

Master of Science
Washington University, 2006

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Epidemiology

The Norman J. Arnold School of Public Health

University of South Carolina

2014

Accepted by:

Wilfried Karmaus, Major Professor

Susan E. Steck, Committee Member

Hongmei Zhang, Committee Member

John W. Holloway, Committee Member

Lacy Ford, Vice Provost and Dean of Graduate Studies

DEDICATION

I dedicate this work to my parents who have provided me with unconditional support throughout my life and who inspired and encouraged me to pursue my educational ambitions. To my two sisters whose support fortified my educational success. A special feeling to my little sister, Fatima, being a Down syndrome girl has always made you so special and you will always have a special place in my heart.

I also dedicate this dissertation to my beloved wife who ensured my success through ongoing support, patience, and encouragement.

To my beloved daughter, Retaj (4 years old), who has been an unmatched source of inspirations during the course of my studies.

To my mentor, Prof. Wilfried Karmaus, who supported me and my family in exceptional manners. Prof. Karmaus, set the bar high from day one and guided me to that bar.

ACKNOWLEDGEMENTS

I am sincerely grateful for the exceptional mentorship that I received from Prof. Wilfried Karmaus, the chair of my dissertation committee, who guided me throughout my PhD studies. Prof. Karmaus capitalized on my strengths and helped me overcome my limitations. I have learned the meaning of 'role model mentorship' through the leadership, dedication, and hard-work of Prof. Karmaus. I am thankful and I will always be for being a student of Prof. Karmaus.

Also, I would like to thank all members of my dissertation committee, Prof. John W. Holloway, Dr. Susan E. Steck, and Dr. Hongmei Zhang, who have provided me with insightful, careful, and constructive comments and suggestions that ensured the success of my dissertation research. I would like to express special gratitude to all of my dissertation committee members who devoted their time and effort, and most of all patience to make this an enjoyable learning process.

I would like thank participants of the Isle of Wight birth cohort and their families who have helped us with this project over the last two decades. I would like to acknowledge the help of all the staff at The David Hide Asthma and Allergy Research Centre in undertaking the 18 year and previous assessments of 1989 Isle of Wight birth cohort.

ABSTRACT

Background: Allergic disorders, including eczema, asthma, and rhinitis, have emerged as a global public health concern due to their elevated prevalence and the associated clinical morbidity. Environmental, immunologic, and genetic factors have been implicated in the pathogenesis of allergic disorders. Allergic sensitization (representing deviated immune responses) and filaggrin gene (*FLG*) variants (leading to dysfunctional epidermal barrier) have shown to be common predisposing factors in the development of allergic disorders. However, there is a lack of knowledge on their joint effects on the development of single and multiple (coexistence) allergic disorders. More recently, epigenetic mechanisms, such as DNA methylation, have emerged as potentially important factors in the development of such complex diseases; however, the extent to which DNA methylation associates with allergic disorders is unclear.

Objectives: This dissertation sought to (i) determine whether eczema and/or allergic sensitization is an effect modifier of the association between '*FLG* variants and asthma' and '*FLG* variants and rhinitis', (ii) test whether *FLG* variants and allergic sensitization jointly predispose to the comorbidity of eczema, asthma and rhinitis, and (iii) examine associations between DNA methylation across the epidermal differentiation complex (EDC) genomic region with eczema status.

Methods: The Isle of Wight (IOW) birth cohort, a population-based sample of 1,456 infants born between January 1989 and February 1990, was prospectively assessed at ages 1, 2, 4, 10, and 18 years. Repeated measurements of eczema, asthma, rhinitis, and

allergic sensitization (documented by skin prick tests) were available for all follow-ups. *FLG* variants R501X, 2282del4, and S3247X were genotyped in 1,150 participants. Log-binomial regression models were applied to test for associations and statistical interactions on multiplicative scale. On the other hand, DNA methylation was measured in a subsample (n = 367) of the IOW participants at age 18 years (*discovery cohort*) and in two semi-independent samples (*replication cohorts I and II*). Associations between eczema status and DNA methylation were assessed using linear regression.

Results: *FLG* variants were associated with increased risk of asthma and rhinitis. Both eczema status ($RR_{interaction}$ = 1.96, $P_{interaction}$ = 0.006) and allergic sensitization ($RR_{interaction}$ = 1.58, $P_{interaction}$ = 0.013) modified the association between *FLG* variants and asthma, but not the association with rhinitis. The combined effect of both risk factors increased the risk of coexisting "eczema and asthma" (RR = 13.67, 95% CI: 7.35 – 25.42), "asthma and rhinitis" (RR = 7.46, 95% CI: 5.07 – 10.98), and "eczema, asthma, and rhinitis" (RR = 23.44, 95% CI: 12.27 – 44.78). On the other hand, Differential DNA methylation of CpG site cg12048339 (located within promoter of *S100A6* gene) was associated with eczema specifically among female participants of all study cohorts; whereas, aberrant DNA methylation of cg10959711 (located within promoter of *S100A11* gene) associated with eczema among male participants in all study samples.

Conclusions: Allergic sensitization and eczema modulated the association between *FLG* variants and asthma, but not rhinitis; implying that the mechanisms and pathways through which *FLG* variants predispose to increased risk of asthma and rhinitis may be different. Moreover, the coexistence of allergic disorders is frequent and allergic sensitization and *FLG* variants jointly increased risk of allergic comorbidities, which may represent more

severe and complex clinical phenotypes. Results of an exploratory investigation demonstrated that DNA methylation of the EDC locus could be an important factor in the development of eczema in a sex-specific manner. Future studies corroborating our findings are needed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Abbreviations

CI.....Confidence Interval

CpG.....Cytosine-Guanine dinucleotides

DNA.....Deoxyribonucleic acid

EDC.....Epidermal Differentiation Complex

FLG.....Filaggrin gene

GEE.....Generalized Estimating Equations

IgE.....Immunoglobulin E

IOW.....Isle of Wight

ISAAC.....International Study of Asthma and Allergies in Childhood

PAF.....Population Attributable Fraction

RR.....Risk Ratio

SA.....Specific Aim

SPT.....Skin Prick Test

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Allergic disorders, including eczema, asthma, and rhinitis, have emerged as a global public health concern [1]. The clinical morbidity associated with allergic disorders that leads to social and economic burden on families and societies further classify these disorders as major global public-health challenges [1, 2]. The direct annual cost associated with clinical management of eczema alone was estimated to be £47 ($76) million in the United Kingdom, with families bearing £17 ($27) million of disease costs [3]. Furthermore, Asthma has even greater economic burden, costing £843 million ($1.368 billion) annually [3]. The elevated economic burden and clinical morbidity associated with such chronic diseases, clearly demonstrates the need for developing cost-effective intervention and prevention strategies that aim at reducing the burden associated with allergic disorders. However, to reach the aforementioned goal, there is a need to better understand the natural history and the biological mechanisms that underlie the clinical manifestation of these complex disorders at different developmental stages.

The International Study of Asthma and Allergies in Childhood (ISAAC) provides unmatched opportunity to explore the trends in prevalence of allergic disorders over time [4]. Reports based on the ISAAC study revealed the existence of global variations in the prevalence of allergic disorders [5-8]. During the last 3 decades, industrialized nations witnessed disproportionally increased trends in the prevalence of eczema, asthma, and

rhinitis [5, 9]. However, recent reports are indicating that a plateau in the prevalence has been reached in most of the countries that experienced increasing trends [10-12]. In contrast, countries with formerly low-prevalence are reporting increasing patterns in the prevalence and incidence of allergic disorders [10-13]. The disparity in the prevalence across nations and the changing trends over time within populations are indicators of the importance of environmental factors, in addition to the genetic elements, in disease pathogenesis. Deeper understanding of the natural history and risk factors of these chronic conditions will help us elucidate their plausible underlying etiological mechanisms.

1.2 Overall Structure of the Dissertation

This document is divided into six chapters, as the following:

i. The first chapter provides background information on the dissertation topic, literature review relevant to the study questions, current gap in knowledge, and the specific aims and hypotheses of this study.

ii. The second chapter explains in details the methodological parts of the study and the statistical analyses that were applied to achieve the goals of this dissertation.

iii. Each of chapter three, four, and five is organized and formatted as a journal research report that have been or will be submitted for publication. Each of these three research-related chapters, also referred to as results chapters, represent one of the three specific aims of this dissertation.

iv. The sixth, last, chapter provides an overall synopsis and discussion of the results of the three research-related chapters and how results of these chapters are synthesized.

1.3 Literature Review

1.3.1 Global Variations in the Prevalence of Allergic Disorders

Reports based on the ISAAC study revealed the existence of global variations and both *between* and *within* countries variations in the prevalence of allergic disorders [5, 10, 11, 14, 15]. For instance, for the age group 6 to 7 years, data based on the ISAAC study showed that the 12-month prevalence of eczema ranged from 0.9% in India to 22.5% in Ecuador [14]. For the age group 13 to 14 years, the prevalence of eczema between countries ranged from 0.2% in China to 24.6% in Columbia [14]. Similarly, the prevalence of asthma symptoms among 13 to 14 years old children ranged from 0.1% in Pune (India) to 16% in Costa Rica [16]. Prevalence of rhinitis showed even wider global variation, which ranged from 1% in India to 45.1% in Paraguay [17]. In addition to estimating the prevalence and time trends of allergic disorders, the ISAAC study allowed researcher to discover wide-range of geographical differences in the prevalence of eczema *within* countries. For example, the ISAAC phase 3 study was conducted in 10 different locations across Spain, with the lifetime prevalence among 6 to 7 years olds being 35.6% in 'A Coruna' and 25.0% in 'Barcelona' [14]. Although the role of environmental factors in explaining the *between* and *within* countries variations in allergic disorders prevalence is essential, there is some strong evidence suggesting that genetic predisposition can potentiate the effects of environmental factors [18]. In

summary, both the *between* and *within* countries differences in the prevalence of allergic disorders are worth remarking as they can contribute to the advancement of etiologic research.

1.3.2 Natural History of Allergic Disorders

Eczema is an inflammatory skin disorder that is characterized by disrupted epidermal barrier function, immunoglobulin E (IgE)-mediated sensitization to food and environmental allergens (i.e., allergic sensitization), and a remitting–relapsing combination of skin dryness and itching [19-21]. Globally, the prevalence of eczema among children and adults has been estimated to be between 15% to 30% and 2% to 10%, respectively, demonstrating wide variation across nations [20, 22]. The three major developmental periods of eczema are: infancy, childhood, and adolescence/adulthood; with 45% of all eczema cases occurring in the first 6 months of life [20, 23]. However, the long-term prognosis (natural history) of eczema is complex due to the relapsing and remitting nature of the disease. Emerging evidence is suggesting that eczema affects more boys during early life stages and becomes more predominant among adolescent girls and thereafter [20, 24]. Recently, we demonstrated that girls develop more eczema during puberty period while more boys outgrow their eczema [25]. Yet, more insights on the etiology of eczema are needed in order for preventable public-health efforts to succeed.

Likewise, population-based studies have shown that the prevalence of both asthma and rhinitis is increasing and disproportionally affecting developed and developing countries [10]. Asthma, accounting for a substantial proportion of

burden/morbidity associated with allergic disorders, is a common chronic disorder affecting up to 25% of children and adults around the world [16, 26-29]. The major pathophysiological hallmarks that contribute to the complex heterogeneity of the clinical manifestations of asthma are recurrent airway inflammation, variable airway obstruction, and bronchial hyper-responsiveness [30, 31]. Similarly, an increase in the prevalence of rhinitis (10% to 30%), which is an inflammatory disease affecting the nasal tissue, has been documented during the last 30 years [32]. Common symptoms of rhinitis are nasal congestion and itching, sneezing, and rhinorrhoea, which negatively impact the quality of life [32, 33]. The natural history of both asthma and rhinitis is influenced by gender [34, 35]. For asthma, a gender reversal occurs around the puberty period; the disorder predominantly affects boys before puberty, whereas, more girls are affected during adolescence and in early adulthood [26, 35-37]. However, the effect of gender on the long-term prognosis of rhinitis is further complicated by allergic sensitization (defined later) [33, 34].

Although it is acceptable to investigate the etiological mechanisms underlying allergic diseases separately; evidence supporting their comorbidity is increasing [38, 39]. Allergic comorbidity (coexistence) is defined as the concurrent presence of more than one allergic disorder (i.e. eczema, asthma, and rhinitis) in the same individual. The elevated comorbidity of asthma (lower-airway disorder) and rhinitis (upper-airways disorder) and their shared pathophysiology further support the concept of "one airway one disease" [12, 40]. In addition to the comorbidity of the allergic respiratory diseases, the concept of "allergic march" suggests that allergic diseases develop in a sequential pattern. Eczema is considered as the first step in the allergic march, usually developing

early in infancy/childhood followed by manifestations of allergic disorders of the airway later in childhood or early adolescence period [41]. The continuum developmental mechanism of allergic diseases is widely discussed in the scientific literature; however, supporting replicable evidence and high consensus is far from complete [41, 42]. Another revealing concept is the "coexistence" of allergic morbidities, which opposes the linear developmental pathway [42-44]. A report based on the German Multicenter Atopy Study (MAS), showed that eczema alone early in life does not support the allergic march, rather the coexistence of eczema and wheezing predicted asthma [45]. Such observation speaks in favor of the coexistence of allergic morbidities instead of a progressive development. The interrelationship between these conditions and the risk factors that either predispose individuals to follow the "allergic march" or "coexistence" remain an open field for research.

1.3.3 Allergic Sensitization and Filaggrin Variants: as Common Risk Factors

A complex interplay between genetic, environmental, and immunological factors is considered to contribute to the pathogenesis of allergic disorders [43, 46]. Allergic sensitization, defined as the genetic susceptibility to produce Immunoglobulin E (IgE) antibodies in response to exposure to environmental or food antigens, is the most common thread linking the manifestations of allergic disorders [46-48]. Arshad *et al* suggested that 30% to 40% of cases of eczema, asthma, and rhinitis in early childhood can be attributed to allergic sensitization [49]. Yet, others suggested that up to 80% of asthmatic patients have concomitants allergic sensitization [50]. Similarly, it is well established that manifestations of eczema and rhinitis are closely associated with allergic

sensitization [23, 32, 46, 48]. Candidate-gene and genome-wide association studies have identified several genetic variants that are associated with the observed variability in IgE levels between allergic and non-allergic individuals [51, 52]. As a result, genetic factors that regulate the immune response dominated the research of allergic diseases for the last 30 to 40 years.

The recent discovery of loss-of-function variants in the filaggrin gene (*FLG*) caused a shift in the research paradigm from mainly focusing on immune related genes to incorporating genetic factors that regulate the formation of the epidermal barrier [53-55]. Filaggrin (filament-aggregating protein) is a key protein for the formation of functional skin barrier that inhibits the penetration of allergens, microbes, and irritants and limits transepidermal water loss [55-57]. Loss-of-function variants within the *FLG* are the strongest and most replicated risk factors for eczema development, yet [58, 59]. Filaggrin haploinsufficiency, i.e. reduction in filaggrin protein expression in heterozygous individuals, is associated with approximately 3-fold increased risk for eczema [60, 61]. It has been estimated that around 25% to 50% of eczema patients are carriers of at least one *FLG* variant, which is associated with an estimated population-attributable risk ranging from 4.2% to 15.2% [58, 61-63]. In addition to its association with eczema, multiple studies showed that *FLG* variants are associated with asthma and rhinitis [56, 60]. One school of thought considers that the association between *FLG* variants and asthma depends on the presence of eczema [59, 64, 65]. However, there is conflicting evidence as to whether the association between *FLG* variants and rhinitis is modified by eczema status [63, 64, 66]. To this end, it is clear that *FLG* variants are significant common thread for the development of the different allergic disorders.

1.3.4 DNA Methylation and the Risk of Eczema

The high heritability of eczema (71% to 84%), asthma (35% to 95%), and rhinitis (33% to 91%) and clustering within families demonstrate the importance of genetic predisposition in their pathogenesis [43, 67, 68]. However, thus far candidate-gene and genome-wide association studies have failed to explain the observed high heritability of allergic diseases. Hence, suggesting that genetic regulatory factors, other than DNA sequence variants, such as epigenetic variants may account for the unexplained genetic effect [69]. Epigenetic regulatory mechanisms are mitotically heritable and can alter gene activity without changing the DNA sequence [70]. DNA methylation, widely studied in epidemiological investigations due to practical and biological reasons [71, 72], along with histone modifications are important epigenetic marks that work hand-in-hand on influencing disease expression [73]. Since eczema is considered as the first manifestation of allergic disorders that can predispose to subsequent allergic morbidities, it is thus important to understand the epigenetic etiologic mechanisms that underlie eczema development. Up to our knowledge, the impact of DNA methylation on the risk of eczema has not been investigated previously. The epidermal differentiation complex (EDC), located on human chromosome 1q21, harbors a dense cluster of genes (including *FLG*) that are involved in the terminal differentiation of keratinocytes that are responsible for the integrity and functionality of the epidermal barrier [74-76]. Previous reports showed that a compromised epidermal barrier predispose to the development of eczema [77, 78]. Therefore, a comprehensive DNA methylation profiling of the EDC genomic region should provide novel insights into the epigenetic contribution in the development of eczema.

1.4 Gaps in Current State of Knowledge

Both, allergic sensitization (deviated immune response) and *FLG* variants (defective epidermal barrier) are important risk factors that link and predispose to the development of allergic disorders[46]. However, the current state of knowledge lacks clear understanding on the joint contribution of these two risk factors. Prior studies focused on eczema as the effect modifier for the association between *FLG* variants and asthma, while ignoring the role of allergic sensitization. Similarly, eczema was investigated as an effect modifier for the association between *FLG* variants and rhinitis, but not allergic sensitization. Furthermore, the single and joint role of *FLG* variants and allergic sensitization on the coexistence of allergic diseases (i.e., eczema, asthma, and rhinitis) has not been previously investigated. On the other hand, no existing knowledge on the role of epigenetic regulatory mechanisms (e.g., DNA methylation) of the EDC genomic loci that controls the epidermal barrier formation and functionality in relation to eczema risk is available.

1.5 Study Proposal, Hypotheses, and Aims

Population-based longitudinal studies provide an unmatched opportunity to investigate the natural history of allergic disorders. Cross-sectional and case-control studies lack the ability of determining the longitudinal effect of risk factors on outcomes. However, follow-up studies with repeated measurements of the outcomes of interest (e.g., eczema, asthma, and rhinitis) and risk factors (e.g., allergic sensitization) facilitate the determination of whether the effects are established early or later in the development of diseases. Furthermore, we can determine whether the effect of time-independent risk

factor (e.g. *FLG* variants) changes over-time. In contrast, for dynamic (time-varying) risk factors (e.g., allergic sensitization) determining their concurrent effects allow the risk to change over-time. This dissertation aims at enhancing our understanding of the role of *FLG* variants, allergic sensitization, and epigenetic modifications on the development of allergic diseases, by examining the following specific hypotheses (**H**):

H1: Allergic sensitization, not eczema, is the main effect modifier of the association between *FLG* variants and asthma.

H2: The association between *FLG* variants and rhinitis is modified by allergic sensitization status, not eczema status.

H3: *FLG* variants and allergic sensitization jointly increase the risk of having single and multiple (coexisting) allergic disorders.

H4: Eczema-affected and eczema-free individuals have different DNA methylation profiles for the EDC genomic region.

Hypothesis 1, 2, and 3 will be tested using data from the Isle of Wight 1989 birth cohort study (n = 1,456). Participants were prospectively followed-up at the ages of 1, 2, 4, 10, and 18 years. Hypothesis 4 will be tested in a subset (n = 367) of Isle of Wight birth cohort participants that had DNA methylation measurements. For the purpose of validation, hypothesis 4 will be further tested in two semi-independent samples (pregnant cohort participant, their partners, and their offspring, the F2 generation). The specific aims (**SA**) of this dissertation are:

SA1: To determine the magnitude of modification imposed by eczema and allergic sensitization on the association between "*FLG* variants and asthma" (**H1**). Similarly, I will determine whether eczema and/or allergic sensitization act as an

effect modifier for the association between "*FLG* variants and rhinitis" (**H2**). Overall, I will test whether the combined (joint) effect of two risk factors (i.e., "*FLG* variants and eczema" or "*FLG* variants and allergic sensitization") yields higher risk of asthma (or rhinitis).

SA2: To test whether *FLG* variants and allergic sensitization jointly predispose to the comorbidity of eczema, asthma and rhinitis (**H3**).

SA3: To associate DNA methylation across the EDC genomic region with eczema status (**H4**).

1.6 How Would Results of this Dissertation Enhance our Knowledge?

Better understanding of the genetic and epigenetic contribution to the development of allergic diseases will be gained by successfully accomplishing the specific aims of this dissertation. We have previously shown that the combined effect of allergic sensitization and *FLG* variants predisposed individuals at higher risk of having concurrent and subsequent eczema [79]. In this dissertation, I will investigate whether allergic sensitization (or eczema) in combination with *FLG* variants increase the risk of having concurrent and subsequent asthma and rhinitis. The current knowledge indicates that *FLG* variants increase the risk of asthma only in the presence of eczema [59, 66]; however, I propose to test whether allergic sensitization is the effect modifier rather than eczema. Furthermore, the current inconsistencies in the associations between *FLG* variants and rhinitis can be attributed to the failure of previous studies in accounting for the possible interaction between *FLG* variants and allergic sensitization on the risk of rhinitis. Therefore, **SA1** of this dissertation will enhance our knowledge by showing that

allergic sensitization is the main effect modifier of the association between *FLG* variants and asthma and *FLG* variants and rhinitis.

Results of **SA1** of this dissertation will show the importance of *FLG* variants and allergic sensitization on the development of single allergic disorders (i.e., asthma and rhinitis). In contrast, **SA2** sought to determine whether *FLG* variants and allergic sensitization are associated with the coexistence of multiple allergic diseases rather than a single morbidity. Previous etiologic research in the field of allergic comorbidities is highly limited and assessments of the joint role of allergic sensitization and *FLG* variants on the development of allergic comorbidities is lacking. Therefore, results of **SA2** will enhance our understanding of the underlying risk factors that predispose individuals at higher risk of having multiple allergic disorders.

In addition to the role of genetic variants in the pathogenesis of allergic diseases, the etiologic role of epigenetic modifications (e.g., DNA methylation) is rapidly emerging and scientifically sound. Recently we showed a significant interaction between *FLG* variants and adjacent differential DNA methylation on the risk of eczema [80]. Results of this report suggested that the association between *FLG* variants and eczema is modulated by DNA methylation. Recent evidence is suggesting that epigenetic mechanisms, such as DNA methylation, that alter gene activity without changing the underlying DNA sequence account for a considerable amount of the unexplained genetic effect found in complex diseases, such as eczema [69, 72]. In this dissertation I propose going beyond the single-gene approach (i.e., *FLG* gene) into incorporating the DNA methylation of the EDC locus at chromosome 1q21 (**SA3**). The EDC includes a dense cluster of approximately 60 genes (including *FLG*) encoding structural and regulatory

proteins that are essential for keratinocyte differentiation [76, 81, 82]. Comprehensively analyzing DNA methylation status of the EDC genomic region will provide novel insights, beyond genetic variants, on the contribution of epigenetic modifications. Therefore, I anticipate that DNA methylation of the EDC will further show the importance of this genomic region for the development of eczema. **SA3** will focus on eczema since it is considered the first manifestation of allergic diseases. The ultimate goal of this dissertation is to improve our ability of characterizing and classifying susceptible individuals, which will lead to an improved clinical management and the development of efficient intervention and prevention strategies.

CHAPTER 2

METHODS

2.1 STUDY DESIGN AND PARTICIPANTS

This dissertation is mainly based on information collected from the Isle of Wight (IOW) birth cohort. An unselected population-based study (n = 1,536) was recruited between January 1989 and February 1990 in the Isle of Wight, United Kingdom, to prospectively study the natural history and etiology of asthma and allergic diseases. The island is close to the British mainland, semi-rural, without heavy industry. Both the Isle of Wight and the study populations are 99% Caucasian. After excluding adoptions, perinatal deaths, and refusals for follow-up, written informed consent was obtained from parents to enrol 1,456 children (95%), with follow-up assessments conducted at ages 1 (n = 1,167), 2 (n = 1,174), 4 (n = 1,218), 10 (n = 1,373), and 18 (n = 1,313) years. Ethics approvals were obtained from the Isle of Wight Local Research Ethics Committee (now named the National Research Ethics Service, NRES Committee South Central – Southampton B) at recruitment and for the 1, 2, 4, 10 and 18 years follow-ups (06/Q1701/34). Detailed questionnaires were completed for each child at each follow-up. When a visit was not possible, a telephone questionnaire was completed or a postal questionnaire sent for completion and return.

For the purpose of testing **H4**, a subsample (*discovery cohort*: n = 367, aged 18 years) of the IOW birth cohort participants were randomly selected for epigenetic study. We also considered two cohorts to test the reproducibility of results obtained from the

discovery cohort. *Replication cohort I* (n = 146; parents of the F2 generation) includes participants of the IOW cohort (28.8% were part of the discovery cohort) plus new participants. Expecting mothers and their partners were assessed before delivery. Replication cohort II (n = 94; F2 generation) includes the F2 newborns of the IOW cohort participants plus new parents. Infants were followed-up at ages 3-, 6-, and 12-months.

2.2 PHENOTYPES

In all assessments of the Isle of Wight birth cohort, eczema was defined as chronic or chronically relapsing, itchy dermatitis lasting more than 6 weeks with characteristic morphology and distribution [83], following Hanifin and Rajka criteria [84]. For asthma, at the 1, 2, and 4-year follow-ups, the medical investigator determined the presence of asthma based on wheeze frequency over the last 12 months and treatment given for asthma or asthma related symptoms. At the 10 and 18 year follow-ups, asthma was defined as having "ever had asthma" and either "wheezing or whistling in the chest in the last 12 months" or "current treatment for asthma", using ISAAC questionnaire [4]. Rhinitis was defined by a positive response to: 'In the past 12 months have you had a problem with sneezing, or a runny or a blocked nose when you did not have a cold or the flu?' [33]. Since the 1-year and 2-year follow-up data on eczema, asthma, and rhinitis were collected in a relatively small time window, we combined them for analytic purposes (reported as 1-or-2 years).

To determine allergic sensitization status, skin prick testing (SPT) at ages 1 and 2 years was performed on children with any symptoms of eczema, asthma, or rhinitis. We combined SPT results for ages 1 and 2 years, since they occurred within a short time

period and will henceforth refer to this as SPT at 1-or-2 years. At 4, 10 and 18 years, regardless of symptoms, SPT was performed on most children attending the research center to a standard battery of common allergens (ALK-Albello, Horsholm, Denmark). Inhalant allergens tested were house dust mite, cat, dog, *Alternaria alternata*, *Cladosporium herbarium*, grass pollen mix, and tree pollen mix. Food allergens tested were cows' milk, soya, hens' egg, peanut and cod. Positive and negative controls were included. Allergic sensitization was defined by having a SPT to at least one allergen test with mean wheal diameter of 3 mm greater than the negative control. Since allergic sensitization is a dynamic rather than a completely stable phenotype, we used the concurrent status and thus allowed the risk to change over time.

2.3 *FLG* GENOTYPING

DNA was extracted from blood or saliva samples from cohort subjects (n = 1,211). Five variants in the *FLG* gene that result in loss of function and are reported to be common in populations of European ancestry were selected for genotyping [79]. DNA samples were interrogated using GoldenGate Genotyping Assays (Illumina, Inc, SanDiego, CA) on the BeadXpressVeracode platform (Illumina, Inc, SanDiego, CA) per Illumina's protocol. In brief, samples were fragmented and hybridized to the pool of allele-specific primer sets. Following an extension/ligation reaction the samples were then hybridized to the Veracode bead pool and processed on the BeadXpress reader. Data were analyzed using the genotyping module of the GenomeStudio Software package (Illumina, Inc, SanDiego, CA). DNA from each subject plus 37 replicate samples were analyzed for a total of 1,248 samples. The quality threshold for allele determination was

set at a GenCall score > 0.25 (scores ≤ 0.25 were "no calls") with n = 1,227 samples (98.3%) retained for further analysis. Analysis of each locus included reclustering of genotyping data using our project data to define genotype cluster positions with additional manual reclustering to maximize both cluster separation and the 50th percentile of the distribution of the GenCall scores across all genotypes (50% GC score). Children were classified as having *FLG* loss-of-function defect if they carry the minor allele for at least one of the following *FLG* null variants: R501X, 2282del, or S3247X. Variants 3702delG and R2447X were not informative in our population due to minor allele frequencies being < 0.1% [79], and thus were not used in defining filaggrin haploinsufficiency.

2.4 DNA Methylation Profiling

DNA was extracted from whole blood collected at age 18 years (*discovery cohort*), whole blood of *replication cohort I* participants', and cord blood of *replication cohort II participants'* [85], and bisulfite-treated for cytosine to thymine conversion using the EZ 96-DNA methylation kit (Zymo Research, CA, USA). Genome-wide DNA methylation was assessed using the Illumina Infinium HumanMethylation450 BeadChip (Illumina, Inc., CA, USA), which interrogates > 484,000 CpG sites associated with approximately 24,000 genes. Arrays were processed using a standard protocol [86]. The BeadChips were scanned using a BeadStation, and the methylation level (beta value) calculated for each queried CpG locus using the Methylation Module of BeadStudio software. After cleaning of the DNA methylation data, beta (β) values presented as the proportion of methylated (M) over methylated (M) and unmethylated (U) sites (β=M/[c+M+U], with c

being constant to prevent dividing by zero) were used to estimate the effect of DNA methylation [87].

2.5 Statistical Analysis

Statistical analyses and approach for each of the three specific aims will be presented separately.

2.5.1 Specific Aim 1: Statistical Analysis

SA1: To determine the magnitude of modification imposed by eczema and allergic sensitization on the association between "*FLG* variants and asthma" (**H1**). Similarly, I will determine whether eczema and/or allergic sensitization act as an effect modifier for the association between "*FLG* variants and rhinitis" (**H2**). Overall, I will test whether the combined (joint) effect of two risk factors (i.e., "*FLG* variants and eczema" or "*FLG* variants and allergic sensitization") yields higher risk of asthma (or rhinitis).

For all of the association analyses that will be performed to approach **SA1**, log-binomial (log linear) regression models will be applied to estimate risk ratios (RRs) and their 95% confidence intervals (95% CIs) using the GENMOD procedure in SAS 9.3 (SAS, Gary, NC, USA). Since the proposed effect modifiers (i.e., allergic sensitization and eczema) and the two outcomes (i.e., asthma and rhinitis) were measured repeatedly at ages 1-or-2, 4, 10, and 18 years, the generalized estimating equations (GEE) method will be used, which accounts for the correlated observations and the within-child effect by employing a covariance matrix solved through an iterative estimating process based on a

working correlation matrix [88]. The appropriate working correlation matrix will be determined by using the Akaike information criterion (AIC).

In an explanatory step, to determine whether the RRs of the association of *FLG* variants with asthma and rhinitis differ according to the presence or absences of the proposed effect modifiers, we evaluated the association in the total study sample and sub-samples based on the status of the possible effect modifiers. Next, interaction terms, on a multiplicative scale, were used to test the additional effect of two co-occurring risk factors on the health outcomes above and beyond their individual effects. Also, to evaluate whether the effect of *FLG* variants across levels of the effect modifiers were statistically significantly different, we included interaction terms in separate regression models (model 1: *FLG* variants × eczema and model 2: *FLG* variants × allergic sensitization). To decipher which interactive effect was more pronounced, a regression model including both interaction terms (model 3: *FLG* variants × eczema and *FLG* variants × allergic sensitization) was evaluated. Henceforth, we refer to models 1, 2, and 3 as the 'concurrent models' since the effect modifiers and the outcomes coexisted at the same time. In the case of a possibly statistically significant interaction term ($P_{\text{interaction}} < 0.1$), the "combined effect", referring to the joint impact of two individual risk factors plus their interaction on the occurrence of the outcome, was estimated as follows: RR = exp[(β1 × allergic sensitization) + (β2 × *FLG* variants) + (β3 × allergic sensitization × *FLG* variants)]. We repeated the previous analyses while stratifying by sex to determine if sex-specific effects existed. In all GEE models, sex and age at follow-up were included as potential confounders.

In addition to the concurrent models, we tested the interactions in 'delayed effect' models to investigate whether the time-order of risk and response supports the concurrent model findings. The delayed effect analysis helps to determine whether the interaction effects contribute to the developmental process of asthma and rhinitis. To this end, we analyzed whether *FLG* variants interact with preceding allergic sensitization or eczema on the risk of subsequent asthma and rhinitis. For example, the interaction of *FLG* variants with allergic sensitization status at age 1-or-2 years on the risk of asthma at age 4 years (*FLG* variants × SPT results at 1-or-2 years → asthma at 4 years) was evaluated. The prevalence of the clinical phenotype (e.g., asthma at 4), which was used in the delayed effect models, included new occurrences and persistence of the disease. Therefore, we constructed three delayed periods (1-or-2 to 4 years, 4 to 10 years, and 10 to 18 years) to determine if the presence of both *FLG* variants and preceding allergic sensitization or preceding eczema influences the positive transition, defined as the change in disease status from disease-free to diseased, and the persistence of the disease in two consecutive follow-ups.

2.5.2 Specific Aim 2: Statistical Analysis

SA2: To test whether *FLG* variants and allergic sensitization jointly predispose to the comorbidity of eczema, asthma and rhinitis (**H3**).

Information on eczema, asthma, and rhinitis were collected prospectively at ages 1-or-2, 4, 10, and 18 years from the IOW study participants. For each age I will create a new variable that has 8 categories, which identifies all the possible combinations of allergic comorbidities at the respective follow-up (Table 2.1). This type of categorization

will allow us to have better understanding on the prevalence of each allergic category at the different ages.

Table 2.1. Categories of allergic comorbidities

No Disease
Rhinitis
Asthma
Eczema
Eczema + Asthma
Eczema + Rhinitis
Asthma + Rhinitis
Eczema + Asthma + Rhinitis

To estimate the single and combined effects of the risk factors, *FLG* variants and repeated measures of allergic sensitization, on allergic disorders, generalized estimating equations (GEE) were used when estimating the covariance matrix to account for correlated observations and the within-child effect of the repeated measurements [88]. Risk ratios (RR) and their 95% confidence intervals (95% CIs) were estimated by applying log-binomial regression models using the GENMOD procedure in SAS 9.3 (SAS, Gary, NC, USA). The "no allergic disorder" category formed the reference category. The main effects of *FLG* variants and allergic sensitization on the risk for single allergic disorders and comorbidities were estimated separately (single risk factor models). In a full model, the two main plus their interaction effect (*FLG* variants × allergic sensitization) were estimated. We repeated the aforementioned analysis while excluding the 1-or-2 years follow-up data from the GEE analysis to estimate whether possible misclassifications of phenotypes at early life biased our results. In the case of a possibly statistically significant interaction ($P_{\text{interaction}} < 0.1$), the combined effect was estimated as follows: RR = exp[(β1 × allergic sensitization) + (β2 × *FLG* variants) + (β3 × allergic sensitization × *FLG* variants)]. The term 'combined effect' was used to

describe the joint impact of two individual risk factors plus their interaction on the occurrence of the outcome. In all GEE models, sex and age at follow-up were included as potential confounders. To evaluate statistical significance of main effects of covariates (not interaction terms), a *P*-value ≤ 0.05 was used as indicator of statistical significance.

To determine the proportion of the different allergic morbidities/comorbidities related to *FLG* variants and allergic sensitization, population attributable fractions (PAFs) were estimated using the following formula: $PAF = p_c \times [(RR - 1)/RR]$, where p_c is the proportion of diseased-individuals that are exposed [89, 90]. This approach yields valid PAF estimates in the presence of confounding and/or effect modification, only when using adjusted RRs [89, 90]. In case of statistically significant interactions between allergic sensitization and *FLG* variants, the RRs for the combined effect were used to estimate PAFs associated with both risk factors. If no interaction was detected, adjusted RRs were used to estimate separate PAFs for each of the two risk factors and an overall PAF was obtained using the following formula: $PAF_{overall} = 1 - [(1 - PAF_{allergic\ sensitization}) \times (1 - PAF_{\textit{FLG}\ variants})]$ [89].

2.5.3 Specific Aim 3: Statistical Analysis

SA3: To associate DNA methylation across the EDC genomic region with eczema status (**H4**).

Although genome-wide DNA methylation profiles were available, I restricted the analysis to 256 CpG sites that span the EDC genomic region (location on human chromosome 1: 151958685 – 153628983). After performing quality control measures on DNA methylation data, *Beta*-values presented as the proportion of methylated (M) over

methylated (M) and unmethylated (U) sites (*Beta*-value = M/[c+M+U], with c being constant to prevent dividing by zero) were calculated for each CpG site [87]. *Beta*-values provide intuitive biological interpretation (i.e., % methylation); however, (i) being constrained between values of 0 (unmethylated) and 1 (completely methylated) and (ii) demonstrating high heteroscedasticity in the lower and upper ends of the methylation range has raised concerns on the validity of this measure when performing statistical analysis [91]. The conversion of *Beta*-values to *M*-values (*M*-value = $\log_2$(*Beta*-value/[1 – *Beta*-value]) has been shown to overcome the limitations of beta-values [91]. Moreover, the use of *M*-values, when performing statistical analysis, demonstrated increased reliability in terms of detecting differentially methylated CpG sites; however, *M*-values cannot be directly interpreted [91, 92]. Therefore, following previous recommendation [91], we used *M*-values when performing the statistical analysis and reported the statistics at the scale of *Beta*-values.

All statistical analyses were conducted using SAS® version 9.3 (SAS Institute, Cary, NC, USA). In the *discovery cohort*, linear regression was used to test the association between DNA methylation (*M*-values) of all CpG sites spanning the EDC region with eczema status adjusting for sex and cell-type composition. To minimize false-positive findings identified in 256 models, CpG sites with a false discovery rate (FDR) adjusted p-value < 0.1 were selected and considered as potentially associated with eczema. The selected CpG sites, in the *discovery cohort*, were taken forward and tested for their possible association with eczema status in *replication cohorts I and II.* Then, the selected CpG sites were tested for association with eczema in males and females separately.

CHAPTER 3

RESULTS I

MANUSCRIPT # 1: ASSOCIATION OF FILAGGRIN VARIANTS WITH ASTHMA AND RHINITIS: IS ECZEMA OR ALLERGIC SENSITIZATION STATUS AN EFFECT MODIFIER?[1]

[1] A. H. Ziyab, W. Karmaus, J. W. Holloway, H. Zhang, S. E. Steck, S. Ewart, S. H. Arshad. Submitted to *International Archives of Allergy and Immunology*, 12/4/2013.

3.1 ABSTRACT

Background: Association of filaggrin (*FLG*) variants with asthma and rhinitis have shown to be modulated by eczema status. However, it is unknown whether allergic sensitization status modifies this association. The aim of this study was to determine whether *FLG* variants need eczema and/or allergic sensitization as a necessary component to execute its adverse effect on coexisting and subsequent asthma and rhinitis.

Methods: Repeated measurements of asthma, rhinitis, eczema, and allergic sensitization (documented by skin prick tests) at ages 1, 2, 4, 10, and 18 years were ascertained in the Isle of Wight birth cohort (n = 1,456). *FLG* haploinsufficiency was defined as having at least the minor allele of R501X, 2282del4, or S3247X variants. Log binomial regression models were used to test associations and statistical interactions.

Results: *FLG* variants increased the risk of asthma (RR = 1.39, 95% CI: 1.06 – 1.80) and rhinitis (RR = 1.37, 95% CI: 1.16 – 1.63). In delayed effect models, '*FLG* variants plus allergic sensitization' and '*FLG* variants plus eczema' increased the risk of subsequent asthma by 4.93-fold (95% CI: 3.61 – 6.71) and 3.33-fold (95% CI: 2.45 – 4.51), respectively, during the first 18 years of life. In contrast, neither eczema nor allergic sensitization in combination with *FLG* variants increased the risk of later rhinitis.

Conclusions: Allergic sensitization and eczema modulated the association between *FLG* variants and asthma, but not rhinitis. Results of our study imply that the mechanisms and pathways through which *FLG* variants predispose to increased risk of asthma and rhinitis may be different.

3.2 INTRODUCTION

The burden associated with allergic disorders, including eczema, asthma, and rhinitis is substantial and of public health importance. Prior investigations indicated that a complex interplay between genetic, immunological, and environmental factors contributes to the development and maintenance of such allergic manifestations. Until recently, etiologic research focused on the role of immune dysregulation in the pathogenesis of allergic disorders; however, a shift in the research paradigm towards understanding the contribution of a defective epidermal barrier rapidly emerged after the discovery of loss-of-function variants in the filaggrin gene (*FLG*) [53, 56, 60]. Filaggrin haploinsufficiency, defined as partial or complete loss of filaggrin (filament-aggregating protein) protein, is associated with the development of an impaired epidermal barrier that is characterized by increased allergen penetration and water permeability [93].

Thus far, *FLG* variants are the most replicated and strongest genetic risk factor for eczema [59]. Also, *FLG* variants are considered to be associated with asthma and rhinitis [66]. Of importance is that the association between *FLG* variants and asthma is stronger in the presence of eczema [59, 66]. However, this possible effect modification by eczema was not demonstrated for the association between *FLG* variants and rhinitis [66]. Mechanisms that underlie the association of *FLG* variants with asthma and rhinitis are not well understood since *FLG* is not expressed in the upper or lower airway epithelium [94, 95]. A proposed pathway is that cutaneous sensitization, facilitated by *FLG* variants, may lead to local and systematic inflammation at distant organs (i.e., lung and nasal tissues) [56, 60].

While the concept that eczema leads to asthma in children who have a loss-of-function variants in the *FLG* gene is attractive, research in this area has neglected the role of allergic sensitization, defined as the propensity to produce immunoglobulin E (IgE) antibodies responses to environmental and food antigens [96]. Only one study based on the German Multicenter Allergy Study (MAS) birth cohort addressed allergic sensitization and reported an interaction between *FLG* variants and food sensitization in the pathogenesis of asthma among children with eczema [97]. Hence, further investigations on whether the associations of *FLG* variants with asthma and rhinitis are modified by eczema and/or allergic sensitization are needed so that preventive efforts can be directed to either clinical management of eczema or allergic sensitization or both.

We hypothesized that allergic sensitization, rather than eczema status, modifies the association of *FLG* variants with asthma and rhinitis. Therefore, statistical interactions (*FLG* variants × allergic sensitization and *FLG* variants × eczema) were tested to determine whether *FLG* variants plus allergic sensitization and/or *FLG* variants plus eczema jointly increased the risk of asthma and/or rhinitis. To this end, analyzing longitudinal data from the Isle of Wight birth cohort covering childhood and adolescence prospectively enabled us to determine whether *FLG* variants need allergic sensitization or eczema as a necessary component to execute its adverse effect on coexisting and subsequent asthma and rhinitis.

3.3 Methods

3.3.1 Study Design and Participants

An unselected whole population birth cohort (n = 1,536) was recruited in 1989 in the Isle of Wight, UK, to prospectively study the natural history of allergic conditions. After exclusion of adoptions, perinatal deaths, and refusal for follow-up, 1,456 (95%) children were enrolled, with follow-up assessments conducted at 1, 2, 4, 10, and 18 years of age. Ethics approvals were obtained from the Isle of Wight Local Research Ethics Committee (NRES Committee South Central – Southampton B) at recruitment and for the subsequent follow-ups (06/Q1701/34). At each follow-up, validated questionnaires, including the International Study of Asthma and Allergy in Childhood (ISAAC) questionnaire [4], were completed on allergic disorders plus demographic attributes and exposures to environmental factors.

3.3.2 Phenotypes

In all assessments of the Isle of Wight birth cohort, eczema was defined as chronic or chronically relapsing, itchy dermatitis lasting more than 6 weeks with characteristic morphology and distribution [83], following Hanifin and Rajka criteria [84]. For asthma, at the 1, 2, and 4-year follow-ups, the medical investigator determined the presence of asthma based on wheeze frequency over the last 12 months and treatment given for asthma or asthma related symptoms. At the 10 and 18 year follow-ups, asthma was defined as having "ever had asthma" and either "wheezing or whistling in the chest in the last 12 months" or "current treatment for asthma", using ISAAC questionnaire [4]. Rhinitis was defined by a positive response to: 'In the past 12 months have you had a

problem with sneezing, or a runny or a blocked nose when you did not have a cold or the flu?' [33] Since the 1-year and 2-year follow-up data on eczema, asthma, and rhinitis were collected in a relatively small time window, we combined them for analytic purposes (reported as 1-or-2 years).

To determine allergic sensitization status, skin prick testing (SPT) at ages 1 and 2 years was performed on children with any symptoms of eczema, asthma, or rhinitis. We combined SPT results for ages 1 and 2 years since they occurred within a short time period and will henceforth refer to this as SPT at 1-or-2 years. At 4, 10, and 18 years, regardless of symptoms, SPT was performed on most children attending the research center to a standard battery of common allergens (ALK-Albello, Horsholm, Denmark). Inhalant allergens tested were house dust mite, cat, dog, *Alternaria alternata*, *Cladosporium herbarium*, grass pollen mix, and tree pollen mix. Food allergens tested were cows' milk, soya, hens' egg, peanut and cod. Positive and negative controls were included. Allergic sensitisation was defined by having a positive SPT to at least one allergen test with mean wheal diameter of 3 mm greater than the negative control. Since allergic sensitization is a dynamic rather than a completely stable phenotype, we used the concurrent and/or preceding status and thus allowed the risk to change over time.

3.3.3 *FLG* GENOTYPING

Blood and/or saliva samples were collected at ages 10 and/or 18 years from which genomic DNA was isolated. DNA samples were interrogated using GoldenGate Genotyping Assays (Illumina, Inc, SanDiego, CA) on the BeadXpressVeracode platform (Illumina, Inc, SanDiego, CA) per Illumina's protocol. Individuals carrying the minor

allele for at least one of the *FLG* variants R501X, 2282del4, or S3247X were classified as having filaggrin haploinsufficiency. The R2447X variant was also genotyped, but none of the study participants carried the minor allele. Detailed information on genotyping is provided by Ziyab *et al.* [79].

3.3.4 Statistical Analysis

Since the proposed effect modifiers (i.e., allergic sensitization and eczema) and the two outcomes (i.e., asthma and rhinitis) were measured repeatedly at ages 1-or-2, 4, 10, and 18 years, we applied the generalized estimating equation (GEE) method with first-order autoregressive covariance matrix to account for the correlated observations and the within-child effect of the repeated measurements [88]. Risk ratios (RR) and their 95% confidence intervals (95% CIs) were estimated by applying log-binomial regression models using the GENMOD procedure in SAS 9.3 (SAS, Gary, NC, USA). To determine whether the RRs of the association of *FLG* variants with asthma and rhinitis differ according to the presence or absences of the proposed effect modifiers, we evaluated the association in the total study sample and sub-samples based on the status of the possible effect modifiers.

Interaction terms, on a multiplicative scale, were used to test the additional effect of two co-occurring risk factors on the health outcomes above and beyond their individual effects. Also, to evaluate whether the effect of *FLG* variants across levels of the effect modifiers were statistically significantly different, we included interaction terms in separate regression models (model 1: *FLG* variants × eczema and model 2: *FLG* variants × allergic sensitization). To decipher which interactive effect was more

pronounced, a regression model including both interaction terms (model 3: *FLG* variants × eczema and *FLG* variants × allergic sensitization) was evaluated. Henceforth, we refer to models 1, 2, and 3 as the 'concurrent models' since the effect modifiers and the outcomes coexisted at the same time. In the case of a possibly statistically significant interaction term ($P_{\text{interaction}} < 0.1$), the "combined effect", referring to the joint impact of two individual risk factors plus their interaction on the occurrence of the outcome, was estimated. We repeated the previous analyses while stratifying by sex to determine if sex-specific effects existed. In all GEE models, sex and age at follow-up were included as potential confounders.

In addition to the concurrent models, we tested the interactions in 'delayed effect' models to investigate whether the time-order of risk and response supports the concurrent model findings. The delayed effect analysis helps to determine whether the interaction effects contribute to the developmental process of asthma and rhinitis. To this end, we analyzed whether *FLG* variants interact with *preceding* allergic sensitization or eczema on the risk of subsequent asthma and rhinitis. For example, the interaction of *FLG* variants with allergic sensitization status at age 1-or-2 years on the risk of asthma at age 4 years (*FLG* variants × SPT results at 1-or-2 years → asthma at 4 years) was evaluated. The prevalence of the clinical phenotype (e.g., asthma at 4), which was used in the delayed effect models, included new occurrences and persistence of the disease. Therefore, we constructed three delayed periods (1-or-2 to 4 years, 4 to 10 years, and 10 to 18 years) to determine if the presence of both *FLG* variants and *preceding* allergic sensitization or *preceding* eczema influences the positive transition, defined as the change

in disease status from disease-free to diseased, and the persistence of the disease in two consecutive follow-ups [37].

3.4 Results

3.4.1 Description of Study Population

Of the 1,456 children enrolled in the study, 1,377 were followed-up at age 1-or-2 years, 1,214 at 4 years, 1,368 at 10 years, and 1,309 at 18 years. The period prevalence ranged between 11.9% to 14.2% for eczema, 14.3% to 17.7% for asthma, and 5.6% to 35.8% for rhinitis (table 3.1). At age 1-or-2 years, SPTs were performed on symptomatic children (n = 515), of which 20.6% had at least one positive SPT response. At all other ages, the majority of participates underwent SPTs with the proportion of positive SPTs being 19.4% at 4 years, 26.9% at 10 years, and 41.4% at 18 years of age. Genotype frequencies of *FLG* variants (table 3.1) were concordant with Hardy-Weinberg equilibrium.

3.4.2 *FLG* Variants and Asthma

In the total study sample, *FLG* variants increased the risk of repeated occurrence of asthma from 1-or-2 to 18 years of age (RR = 1.39, 95% CI: 1.06 – 1.80; figure 3.1). In the presence of eczema, the effect size increased (RR = 1.56, 95% CI: 1.10 – 2.21); in the absence of eczema, the magnitude of the association between *FLG* variants and asthma was reduced and lost statistical significance (RR = 1.26, 95% CI: 0.93 – 1.69). This association was noticeably modified when stratifying based on the presence (RR = 1.49,

95% CI: 1.14 – 1.94) or absence (RR = 0.76, 95% CI: 0.48 – 1.21) of allergic sensitization.

To determine whether the observed differences in the associations were statistically significant, statistical interactions on a multiplicative scale were evaluated in concurrent models. There was no evidence for an interaction between *FLG* variants and eczema on the risk for asthma (interaction effect RR = 1.18, 95% CI: 0.78 – 1.79, $P_{\text{interaction}} = 0.429$; figure 3.1). However, a statistically significant interaction was found between *FLG* variants and allergic sensitization on the risk for asthma (interaction effect RR = 2.00, 95% CI: 1.22 – 3.28, $P_{\text{interaction}} = 0.006$; figure 3.1). The combined effect of *FLG* variants and allergic sensitization increased the risk of asthma by 3.63-fold (95% CI: 2.81 – 4.70). Furthermore, to determine which interactive effect is more important, we simultaneously included both interaction terms in one regression model. Results showed that the interaction between *FLG* variants and allergic sensitization remained statistically significant after adjusting for the interactive effect between *FLG* variants and eczema (data not shown). However, the latter did not gain statistical significance.

We retested the results of the aforementioned concurrent models, using delayed effect models. In these models we used the *preceding* cumulative SPTs results (or eczema status), i.e. all SPTs results (or eczema status) from assessments that are prior to the outcome assessment, when testing the interaction between *FLG* variants and allergic sensitization (or eczema) on the risk of subsequent asthma. The results of the delayed effect models show that the combined effect of *preceding* allergic sensitization with *FLG* variants increased the risk of asthma in the next exam at each delay period (figure 3.2a). For instance, the presence of both *FLG* variants plus positive SPT at 1-or-2 and/or 4 years

increased the risk of asthma at 10 years (RR = 7.22, 95% CI: 4.84 – 10.80). Although results of the concurrent model (figure 3.1) did not show an interaction between *FLG* variants and eczema, the delayed effect models demonstrated interactive effects between *FLG* variants and *preceding* eczema on the risk of subsequent asthma (figure 3.2b). The repeated measurement analysis shows that the combined effect of *FLG* variants and *preceding* eczema, on average, increased the risk of later asthma during the first 18 years of life (RR = 3.33, 95% CI: 2.45 – 4.51).

To further understand whether the observed interactions between *FLG* variants and preceding allergic sensitization and eczema influence new occurrence and/or persistence of asthma, we calculated the proportions of positive transitions (defined as the change in disease status from disease-free to diseased in two consecutive follow-ups) and persistence among those with and without both risk factors (table 3.2). The proportion of positive transition was more frequent among those with *FLG* variants and allergic sensitization as compared to those without both risk factors across all transition periods. For instance, in the presence of both *FLG* variants and allergic sensitization, 5/16 (31.3%) of those without asthma at age 4 years developed asthma at age 10 years; whereas, only 23/532 (4.3%) of those without both risk factors developed asthma at 10 years. Furthermore, the presence of both risk factors influenced the persistence of asthma (table 3.2). For example, 18/21 (85.7%) of participants with both risk factors who had asthma at ages 10 years continued to have asthma at 18 years, as compared to 25/44 (56.8%) among those without the two risk factors. Similarly, the presence of both *FLG* variants and preceding eczema influenced the positive transition and persistence of asthma.

3.4.3 *FLG* Variants and Rhinitis

FLG variants were associated with increased risk for rhinitis from age 1-or-2 to 18 years in the total study sample of repeated measurements (RR = 1.37, 95% CI: 1.16 – 1.63; figure 3.3). In both the presence (RR = 1.57, 95% CI: 1.23 – 2.00) and absence (RR = 1.28, 95% CI: 1.04 – 1.57) of eczema, the association between *FLG* variants and rhinitis was statistically significant. However, we observed a statistically significant association between *FLG* variants and rhinitis only in the presence of allergic sensitization (RR = 1.34, 95% CI: 1.17 – 1.53), and not in its absence (RR = 1.13, 95% CI: 0.82 – 1.55; figure 3.3).

To test if the observed heterogeneous effects across the levels of the potential effect modifiers are statistically different, we tested multiplicative statistical interactions in concurrent models. There was no evidence for an interactive effect neither between *FLG* variants and eczema (interaction effect RR = 1.15, 95% CI: 0.83 – 1.60, $P_{\text{interaction}}$ = 0.409; figure 3.3) nor between *FLG* variants and allergic sensitization (interaction effect RR = 1.07, 95% CI: 0.76 – 1.51, $P_{\text{interaction}}$ = 0.698; figure 3.3) on the risk for rhinitis. Concordant with results of concurrent models, the delayed effect models did not show any interactions between *FLG* variants and *preceding* allergic sensitization or eczema on the risk of subsequent rhinitis (figure 3.4).

3.4.4 Sex Differences and Time Trends

Separate analyses for each sex were performed and the results for boys and girls were similar (data not shown) and in agreement with the results of analyzing both sexes together in the repeated measurement (GEE) analysis. Furthermore, at each follow-up we

analyzed boys and girls separately to determine whether the effects differ across both time and sex. The obtained results indicated that neither age of exam nor sex influenced our results.

3.5 Discussion

The objective of this study was to determine whether eczema and/or allergic sensitization status act as effect modifiers for the association of *FLG* variants with asthma and rhinitis. In this study, *FLG* variants increased the risk of asthma and rhinitis during the first 18 years of life in the total study sample. In concurrent effect models, interaction between *FLG* variants and allergic sensitization resulted in a combined effect that increased the risk of coexisting asthma by 3.63-fold. Delayed effect models, which take the time order of risk factors and disease occurrences into account, supported the findings of the concurrent effect models regarding allergic sensitization and *FLG* variants. In addition, the delayed effect models showed that *FLG* variants interact with preceding eczema status on the development of subsequent asthma, indicating that both allergic sensitization and eczema act as effect modifiers for the association between *FLG* variants and asthma. In contrast, neither eczema nor allergic sensitization modified the association of *FLG* variants with rhinitis in a statistically significant manner.

The association between *FLG* variants and asthma among those with eczema was first reported by Palmer *et al.* [53]. Subsequently, several studies have replicated this association and added a possible association between *FLG* variants and rhinitis [59, 66]. Results of our study further support the general agreement that *FLG* variants are associated with asthma in the presence of eczema. We further demonstrated that both

preceding allergic sensitization and eczema modified the association between *FLG* variant and asthma through multiplicative interaction (figure 3.2). These findings indicate that *FLG* variants need eczema or allergic sensitization to execute their adverse effects on asthma. The finding of an interaction between *FLG* variants and allergic sensitization in this study improves our understanding of the possible link between genetics of the epidermal barrier (i.e., *FLG* variants) and a respiratory disorder (i.e., asthma). Such an observation further supports the hypothesis that cutaneous sensitization priming, facilitated by *FLG* variants, may migrate to the airways and cause local and systematic inflammation [98]. Although prior studies have widely suggested such a pathway, the majority of previous investigations did not take the status of allergic sensitization into account. An exception is the study by Marenholz *et al.*, which showed an interaction between *FLG* variants and food sensitization on the risk of asthma in children with eczema [97], however, their analytical sample was restricted to children with eczema.

In the delayed effect models we used the prevalence of asthma as the outcome, which does not distinguish between new occurrences and persistence of asthma. Therefore, in additional descriptive analysis, we demonstrated that the presence of both *FLG* variants and preceding allergic sensitization or preceding eczema influenced the new occurrences (positive transition) and persistence of asthma. Hence, our data suggest that the presence of both risk factors ('*FLG* variants and preceding allergic sensitization' or '*FLG* variants and preceding eczema') plays an important role in the development and persistence of asthma.

Attempts using the composite of 'atopic asthma' phenotype were conducted to investigate whether allergic sensitization is the link between *FLG* variants and asthma.

Inconsistent results for the association of *FLG* variants with 'atopic asthma' have been reported [62, 64, 99, 100]. These inconsistencies could be attributed to the use of the composite 'atopic asthma' phenotype, which does not clearly define the reference (no risk) group. By taking this approach investigators tend to group 'non-atopic asthmatics' and 'atopics without asthma' participants together in the reference group. Therefore, when assessing the association of *FLG* variants with 'atopic asthma' one should be aware that the reference group could include atopic participants, which could result in distorting assessments. The advantage of our approach, using two separate risk factors, over the composite 'atopic asthma' phenotyping, is that we were able to estimate the additional effect due to interaction (*FLG* variants × allergic sensitization) that is above and beyond their independent effects. Hence, allowing us to determine whether the combined effect of both risk factors is more than just the multiplication of their independent effects.

There is conflicting evidence in the literature as to whether the association between *FLG* variants and rhinitis is modulated by eczema status. The majority of previous studies found a stronger association between *FLG* variants and rhinitis in the presence of coexisting eczema when compared to their association in the absence of concurrent eczema [62, 63, 65, 66, 101]. Another investigation reported an association between *FLG* variants and 'persistent allergic rhinitis' phenotype [99]. Controversially, however, a significant association has also been reported between *FLG* variants and the composite 'allergic rhinitis' phenotype adjusted for eczema status [64]. Our analyses did not reveal interaction between eczema ($P_{\text{interaction}} = 0.409$) or allergic sensitization ($P_{\text{interaction}} = 0.698$) with *FLG* variants on the risk of rhinitis in the repeated measurement analysis. Our results of an interactive effect between *FLG* variants and allergic

sensitization on the risk of asthma and the lack of such interaction on the risk of rhinitis suggest that the pathway through which *FLG* variants predispose to asthma and rhinitis may be different. For instance, the expression of filaggrin protein was not detected in the human bronchial epithelium; however, it has been shown that filaggrin is expressed in the nasal vestibule [64, 94]. That said, the extent to which filaggrin expression in the nasal vestibular lining influence the development of rhinitis is yet to be investigated by future studies.

Major strengths of our prospective, 18-year study are the repeated phenotyping, objective assessments of allergic sensitization status, and the low loss to follow-up (ranged from 5% to 17%). Moreover, the majority (80%) of the study participants were genotyped for *FLG* variants common among populations of European ancestry [57]. We showed, previously, that the genotyped study participants did not differ from the total cohort with regard to multiple characteristics [79]. Hence, there is no indication of selection bias that could pose a threat to the validity of our study. Misclassification of eczema cases is minimal since a high proportion of subjects showed typical manifestation of eczema in the usual locations (antecubital or popliteal fossae, ankles, face or neck for 97% at 1 year, 91% at 2 years, 75% at 4 years, 86% at 10 years and 76% at 18 years) [25].

Potential limitations are the definition of asthma and rhinitis symptoms in early life. Our asthma conclusion at ages 10 and 18 years followed the ISAAC criteria [4], which was at that time not available for assessments at age 1, 2, and 4 years. Although slightly different methods were used to define asthma, in a previous report we have shown that the minor change in asthma definition over time did not influence the validity

of our asthma classification [37]. In addition, since it is difficult to differentiate between infectious and other forms of rhinitis in infancy, the elevated prevalence at 1-or-2 years might have been influenced by misclassifying 'viral induced infectious rhinitis' [33]. However, by applying repeated measurements analysis including assessments later in childhood and adolescence, we believe that the influence of the possible misclassification of rhinitis at 1-or-2 years on the overall results of the study is minimal, since similar results were obtained when we reran our analyses excluding the 1-or-2 year data. Moreover, since SPTs at age 1-or-2 years were performed on symptomatic children, we speculate that selection bias leading to underestimated RRs might have been induced. Although proportions of SPT positivity at ages 1-or-2 years (20.6%) and 4 years (19.6%; SPT performed irrespective of symptoms) were similar, RRs associating allergic sensitization at age 1-or-2 years with outcomes (i.e., eczema, asthma, and rhinitis) at the same age tended to be smaller than RRs relating allergic sensitization at age 4 years with outcomes at the same age (data not shown), since also the reference group had some symptoms. Hence, the estimated RRs at age 1-or-2 might be underestimated; however, the extent of the possible selection bias is minimal since our results did not noticeably change when we excluded the 1-or-2 years follow-up data from the analyses. Such observations further demonstrate the robustness of the GEE method in providing population averaged estimates (RRs) while accounting for the within-subject correlations of repeated measurements. The use of a slightly relaxed statistical significance threshold ($P_{\text{Interaction}} < 0.1$) for detecting interaction-effects is further supported by the fact that majority of epidemiologic studies lack sufficient power to detect higher order-terms [102].

In conclusion, our results indicate that allergic sensitization and eczema status are effect modifiers of the association between *FLG* variants and asthma. The combined effects of 'preceding allergic sensitization and *FLG* variants' and 'preceding eczema and *FLG* variants' increased the risk of subsequent asthma by 4.93-fold and 3.33-fold, respectively, during the first 18 years of life. These findings suggest that *FLG* variants need allergic sensitization or eczema to execute its adverse effects on asthma. In contrast, neither eczema nor allergic sensitization status statistically significantly modulated the association between *FLG* variants and rhinitis. Hence, results of our study suggest that the mechanisms and pathways through which *FLG* variants, representing impaired epidermal barrier, predispose to increased risk of asthma and rhinitis may be different. Future studies confirming our reported observations and exploring the differential etiological pathways underlying the development of asthma and rhinitis are needed to better identify and stratify those who share similar risk characteristics.

3.6 Acknowledgments

We thank participants and their families who have helped us with this project over the last two decades. We would like to acknowledge the help of all the staff at The David Hide Asthma and Allergy Research Centre in undertaking the 18 year and previous assessments of 1989 Isle of Wight birth cohort. The 18 year assessment of the IOW birth cohort and future analyses were funded by grants from the National Institute of Health, USA (R01 HL082925 and R01AI091905).

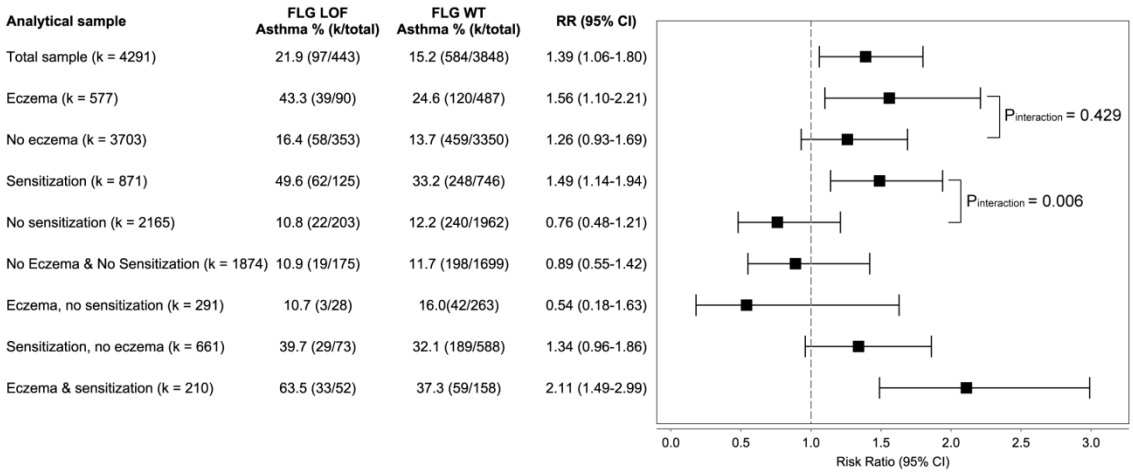

Figure 3.1. Concurrent association of *FLG* variants with asthma in total study sample and sub-samples: longitudinal analysis covering 1-or-2 to 18 years of age. The prevalence of asthma among those with filaggrin loss-of-function (*FLG* LOF) variants and those with wild-type genotype (*FLG* WT) is shown. RR: Risk ratio; 95% CI: 95% confidence interval; k= number of repeated measurements.

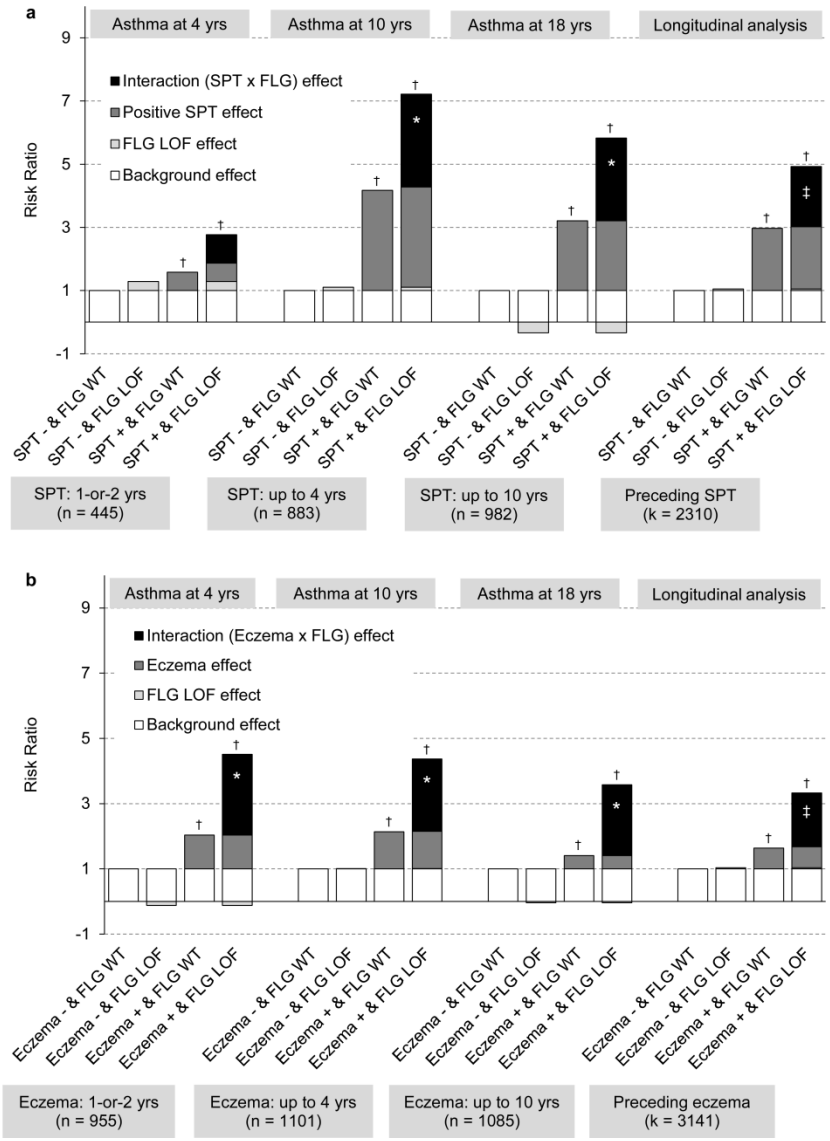

Figure. 3.2. a) Delayed effect models exploring the individual and combined effects of *FLG* variants and preceding allergic sensitization status on asthma development. Independent effects of *FLG* variants (SPT - & *FLG* LOF), positive SPT (SPT + & *FLG* WT), and their combined effect (SPT + & *FLG* LOF) in the development of asthma are shown for different ages. We modeled the effect of preceding SPT results on later asthma. **b)** Delayed effect models exploring the individual and combined effects of *FLG* variants and preceding eczema status on asthma development. Independent effects of *FLG* variants (Eczema - & *FLG* LOF), eczema (Eczema + & *FLG* WT), and their combined effect (Eczema + & *FLG* LOF) in the development of asthma are shown for different ages. We modeled the effect of preceding eczema results on later asthma. SPT +: positive SPT, SPT -: negative SPT; Eczema +: positive eczema diagnosis, Eczema -: negative eczema diagnosis; *FLG* LOF: filaggrin loss-of-function; *FLG* WT: filaggrin wild-type genotype; n = number of children; k= number of repeated measurements. The asterisk (*) and the double dagger (‡) indicate that the interaction effect (**a**: SPT × *FLG*; **b**: Eczema × *FLG*) is possibly statistically significantly different from the null value of 1 (*: $P_{Interaction} < 0.1$; ‡: $P_{Interaction} < 0.05$); suggesting interaction on multiplicative scale. The dagger (†) indicates that the risk estimate (RR) of the given column is statistically significantly ($P < 0.05$) different from the column without both risk factors (**a**: SPT - & *FLG* WT; **b**: Eczema & *FLG* WT). The presence of both an asterisk (*) or double dagger (‡) and dagger (†) signifies that the combined effect of the two risk factors is above and beyond their individuals effects.

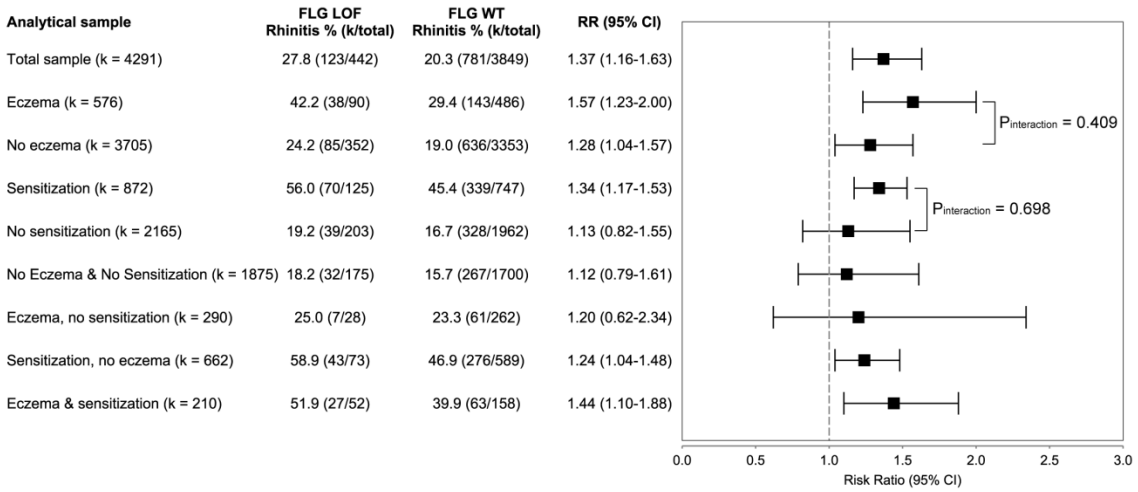

Figure 3.3. Concurrent association of *FLG* variants with rhinitis in total study sample and sub-samples: longitudinal analysis covering 1-or-2 to 18 years of age. The prevalence of rhinitis among those with filaggrin loss-of-function (*FLG* LOF) variants and those with wild-type genotype (*FLG* WT) is shown. RR: Risk ratio; 95% CI: 95% confidence interval; k= number of repeated measurements.

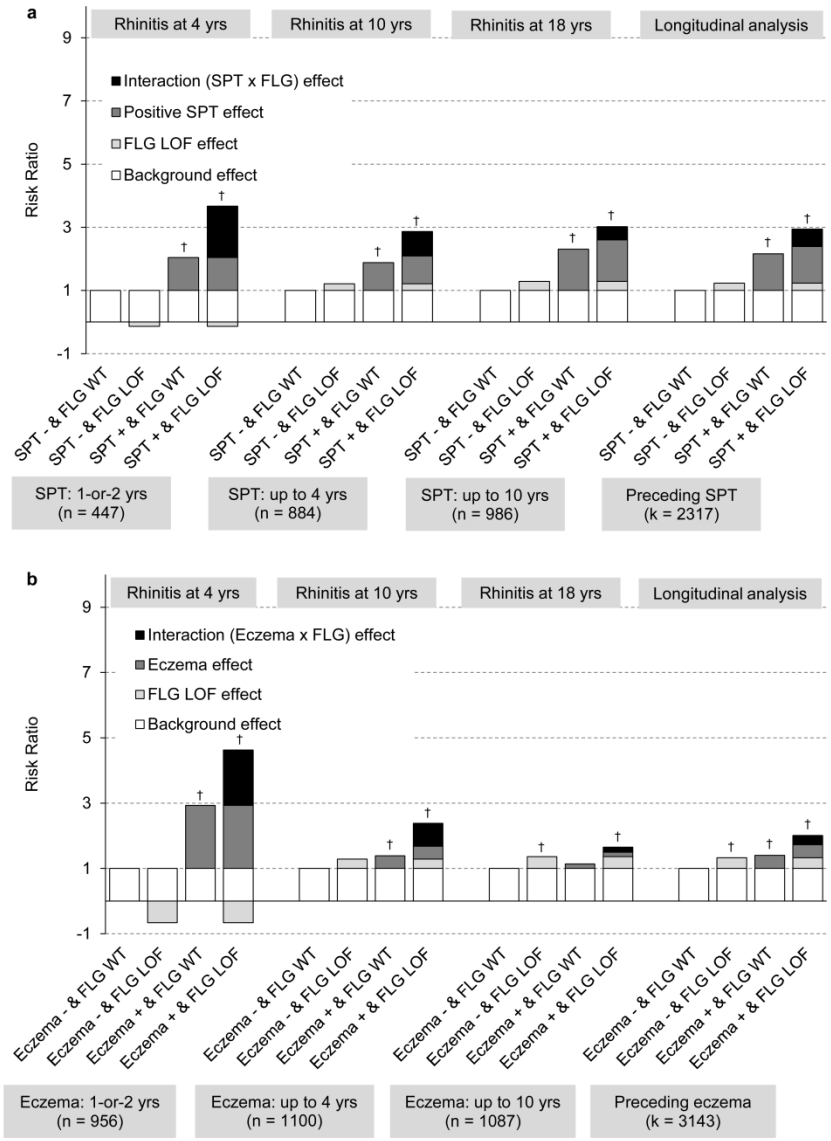

Figure 3.4. a) Delayed effect models exploring the individual and combined effects of *FLG* variants and preceding allergic sensitization status on rhinitis development. Independent effects of *FLG* variants (SPT - & *FLG* LOF), positive SPT (SPT + & *FLG* WT), and their combined effect (SPT + & *FLG* LOF) in the development of rhinitis are shown for different ages. We modeled the effect of preceding SPT results on later rhinitis. **b)** Delayed effect models exploring the individual and combined effects of *FLG* variants and preceding eczema status on rhinitis development. Independent effects of *FLG* variants (Eczema - & *FLG* LOF), eczema (Eczema + & *FLG* WT), and their combined effect (Eczema + & FLG LOF) in the development of rhinitis are shown for different ages. We modeled the effect of preceding eczema results on later rhinitis. SPT +: positive SPT, SPT -: negative SPT; Eczema +: positive eczema diagnosis, Eczema -: negative eczema diagnosis; *FLG* LOF: filaggrin loss-of-function; *FLG* WT: filaggrin wild-type genotype; n = number of children; k= number of repeated measurements. The dagger (†) indicates that the risk estimate (RR) of the given column is statistically significantly ($P < 0.05$) different from the column without both risk factors (**a**: SPT - & *FLG* WT; **b**: Eczema & *FLG* WT).

Table 3.1. Characteristics of study population

Attributes	**% (n/total)**
Sex	
Male	51.2 (786/1536)
Eczema at	
1-or-2 years	14.2 (196/1377)
4 years	11.9 (145/1214)
10 years	13.7 (186/1359)
18 years	12.3 (161/1307)
Asthma at	
1-or-2 years	14.3 (197/1377)
4 years	14.9 (181/1214)
10 years	14.7 (201/1368)
18 years	17.7 (231/1305)
Rhinitis at	
1-or-2 years	15.8 (217/1377)
4 years	5.6 (65/1214)
10 years	22.6 (308/1362)
18 years	35.8 (468/1309)
FLG variants[*]	
R501X	4.1 (47/1161)
2282del4	4.6 (54/1168)
S3247X	1.6 (18/1165)
Combined	10.3 (118/1150)
Allergic sensitization[†] at	
1-or-2 years	20.6 (106/515)
4 years	19.6 (192/982)
10 years	26.9 (279/1036)
18 years	41.4 (353/853)

[*] Analyses were conducted using the combined carrier frequency of 10.3%.

[†] Presence of allergic sensitization was defined by having on or more positive skin prick test result at the respective assessment.

Table 3.2. Prevalence, positive transition, and persistence of asthma stratified based on *FLG* variants and *preceding* allergic sensitization and eczema status: comparing proportions of those with both risk factors to those without the risk factors

	Asthma % (n/total)		
Prevalence	**Age 4 years**	**Age 10 years**	**Age 18 years**
FLG WT & SPT –	18.4 (57/310)	7.9 (47/595)	10.3 (62/604)
FLG LOF & SPT +	50.0 (9/18)	60.0 (18/30)	54.8 (23/42)
p-value*	0.003	< 0.001	< 0.001
FLG WT & Eczema –	11.9 (86/725)	11.2 (91/811)	16.0 (118/739)
FLG LOF & Eczema +	52.0 (5/6)	51.5 (17/33)	51.3 (20/39)
p-value*	< 0.001	< 0.001	< 0.001
Positive transition†	**1-or-2 to 4 years**	**4 to 10 years**	**10 to 18 years**
FLG WT & SPT –	8.5 (19/223)	4.3 (23/532)	6.7 (37/550)
FLG LOF & SPT +	50.0 (6/12)	31.3 (5/16)	23.8 (5/21)
p-value*	< 0.001	< 0.001	0.014
FLG WT & Eczema –	8.5 (53/635)	7.0 (44/632)	9.3 (59/639)
FLG LOF & Eczema +	42.1 (8/19))	29.4 (5/17)	14.3 (3/21)
p-value*	< 0.001	0.006	0.437
Persistence‡	**1-or-2 to 4 years**	**4 to 10 years**	**10 to 18 years**
FLG WT & SPT –	43.7 (38/87)	38.7 (24/62)	56.8 (25/44)
FLG LOF & SPT +	50.0 (3/6)	92.9 (13/14)	85.7 (18/21)
p-value*	0.542	< 0.001	0.026
FLG WT & Eczema –	36.7 (33/90)	44.7 (38/85)	68.3 (56/82)
FLG LOF & Eczema +	83.3 (5/6)	80.0 (12/15)	94.4 (17/18)
p-value*	0.031	0.023	0.037

FLG LOF: filaggrin loss-of-function; *FLG* WT: filaggrin wild-type genotype; SPT -: negative skin prick test result; SPT +: positive skin prick test result.

* *p*-values comparing the proportions across the two exposure groups were derived from fisher's exact test.

† Positive transition refers to the change in disease status from asthma-free to asthma in two consecutive assessments.

‡ Persistence refers to the proportion of individuals who had asthma at two consecutive assessments

CHAPTER 4

RESULTS II

MANUSCRIPT # 2: ALLERGIC SENSITIZATION AND FILAGGRIN VARIANTS PREDISPOSE TO THE COMORBIDITY OF ECZEMA, ASTHMA, AND RHINITIS: RESULTS FROM THE ISLE OF WIGHT BIRTH COHORT[2]

[2] A. H. Ziyab, W. Karmaus, J. W. Holloway, H. Zhang, S. E. Steck, S. Ewart, S. H. Arshad. *Clinical & Experimental Allergy*. doi: 10.1111/cea.12321. Reprinted here with permission of publisher.

4.1 ABSTRACT

Background: Allergic sensitization and filaggrin gene (*FLG*) variants are important risk factors for allergic disorders; however, knowledge on their individual and interactive effects on the coexistence of eczema, asthma, and rhinitis is lacking.

Objective: This study aimed at investigating the single and combined effects of allergic sensitization and *FLG* variants on the development of single and multiple allergic disorders.

Methods: The Isle of Wight Birth Cohort (n = 1,456) has been examined at 1, 2, 4, 10, and 18 years of age. Repeated measurements of eczema, asthma, rhinitis, and skin prick tests were available for all follow-ups. *FLG* variants were genotyped in 1,150 participants. Associations of allergic sensitization and *FLG* variants with single and multiple allergic disorders were tested in log-binomial regression analysis.

Results: The prevalence of eczema-, asthma-, and rhinitis-only ranged from 5.6% to 8.5%, 4.9% to 10.2%, and 2.5% to 20.4%, respectively, during the first 18 years of life. The coexistence of allergic disorders is common, with approximately 2% of the population reporting the comorbidity of "eczema, asthma, and rhinitis" during the study period. In repeated measurement analyses, allergic sensitization and *FLG* variants, when analyzed separately, were associated with having single and multiple allergic disorders. Of particular significance, their combined effect increased the risk of "eczema and asthma" (RR = 13.67, 95% CI: 7.35 – 25.42), "asthma and rhinitis" (RR = 7.46, 95% CI: 5.07 – 10.98), and "eczema, asthma, and rhinitis" (RR = 23.44, 95% CI: 12.27 – 44.78).

Conclusions and Clinical Relevance: The coexistence of allergic disorders is frequent and allergic sensitization and *FLG* variants jointly increased risk of allergic

comorbidities, which may represent more severe and complex clinical phenotypes. The interactive effect and the elevated proportion of allergic comorbidities associated with allergic sensitization and *FLG* variants emphasize their joint importance in the pathogenesis of allergic disorders.

4.2 Introduction

Allergic disorders, including eczema, asthma, and rhinitis, pose social and economic burden on individuals, families, and societies [1, 2]. Worldwide, the lifetime prevalence of eczema symptoms (e.g., itchy rash) is estimated to be 15% to 30% among children and 2% to 10% among adults [19]. Similarly, the prevalence of asthma (up to 25%) and rhinitis (up to 30%) is high [16, 29, 32]. Natural history investigations have demonstrated close and complex relationships between these allergic disorders [25, 33, 48, 103-105]. However, mechanisms and pathways underlying their development is an area of ongoing scientific dispute.

The "allergic march" concept suggests that allergic disorders develop in a sequential pattern, starting with eczema in early infancy and later in childhood developing asthma and rhinitis [106]. However, supporting replicable evidence and consensus is far from complete [42, 107, 108]. An opposing concept is the "coexistence" of allergic morbidities [42, 45, 109]. A report based on the German Multicenter Allergy Study showed that single occurrence of eczema early in life does not result in an allergic march, but the coexistence of eczema and wheezing predicted asthma [45]. Results from the BAMSE birth cohort showed that the comorbidity of allergic disorders is frequent during the first 12 years of life [109]. Such observations speak in favor of coexistence of

allergic morbidities instead of a progressive development. The interrelationship between these conditions and the risk factors that predispose individuals to develop multiple allergic disorders is an open field for research.

Interplay between genetic, environmental, and immunological factors is considered to contribute to the pathogenesis of allergic disorders [110]. Allergic sensitization, defined as the susceptibility to produce immunoglobulin E (IgE) antibodies in response to antigens, is widely-considered as a common thread linking various manifestations of allergic disorders [49, 96]. Likewise, loss-of-function variants in the filaggrin gene (*FLG*) resulting in an impaired epidermal barrier have rapidly emerged as common risk factor for the development of allergic disorders [60]. However, there is a gap in understanding of the respective contribution of *FLG* variants and allergic sensitization on the development of allergic comorbidities. Moreover, the joint role of *FLG* variants and allergic sensitization on the coexistence of eczema, asthma, and rhinitis has not been previously investigated. Using data from the Isle of Wight (IOW) birth cohort, we aimed at determining the single and combined effects of *FLG* variants and allergic sensitization on the development of single and multiple (coexisting) allergic disorders.

4.3 Materials and Methods

4.3.1 Study Design and Participants

An unselected whole population birth cohort (n = 1,536) was recruited in 1989 on the Isle of Wight, UK, to prospectively study the natural history of allergic conditions. After exclusion of adoptions, perinatal deaths, and refusal for follow-up, written informed

consent was obtained from parents to enroll 1,456 (95%) newborns, with follow-up assessments conducted at 1, 2, 4, 10, and 18 years of age. Ethics approvals were obtained from the Isle of Wight Local Research Ethics Committee (now named the National Research Ethics Service, NRES Committee South Central – Southampton B) at recruitment and for the subsequent follow-ups (06/Q1701/34).

4.3.2 PHENOTYPES

In all assessments of the IOW birth cohort, eczema was defined as chronic or chronically relapsing, itchy dermatitis lasting more than 6 weeks with characteristic morphology and distribution [83], following Hanifin and Rajka criteria [84]. For asthma, at the 1, 2, and 4-year follow-ups, the medical investigator determined the presence of asthma based on wheeze over the last 12 months and treatment given for asthma or asthma related symptoms. At the 10 and 18 year follow-ups, asthma was defined as having "ever had asthma" and either "wheezing or whistling in the chest in the last 12 months" or "current treatment for asthma", following the International Study of Asthma and Allergies in Childhood (ISAAC) criteria [4]. Rhinitis was defined by a problem with sneezing, or a runny or a blocked nose without a cold or the flu in the last 12 months [33]. Since the 1-year and 2-year follow-up data were collected in a relatively small time window, we combined them for analytic purposes (reported as 1-or-2 years).

To determine allergic sensitization, skin prick tests (SPTs) at ages 1 and 2 years were performed on children with any symptoms of eczema, asthma, or rhinitis. Parallel to the disorders, we combined the SPTs at ages 1 and 2 years (reported as 1-or-2 years). At 4, 10, and 18 years, regardless of symptoms, SPTs were performed on most children

attending the research center using 14 common food and aeroallergens (ALK-Albello, Horsholm, Denmark) [111]. Allergic sensitization was defined by having a positive SPT to at least one test with mean wheal diameter of 3 mm greater than the negative control. Since allergic sensitization is a dynamic phenotype, we used the concurrent status and thus allowed the risk to change over time.

4.3.3 *FLG* GENOTYPING

Blood and/or saliva samples were collected at ages 10 and/or 18 years from which genomic DNA was isolated. DNA samples were interrogated using GoldenGate Genotyping Assays on the BeadXpressVeracode platform (Illumina, Inc, SanDiego, CA) per Illumina's protocol. Individuals carrying the minor allele for at least one of the *FLG* variants R501X, 2282del4, or S3247X were classified as having filaggrin haploinsufficiency. Detailed information on genotyping is provided by Ziyab *et al* [79].

4.3.4 STATISTICAL ANALYSIS

The 12-month period prevalence of eczema, asthma, rhinitis, and their comorbidities combinations was determined for each follow-up. Combinations of three disorders resulted in eight non-overlapping comorbidities groups (Table 4.1). To estimate the single and combined effects of the risk factors, *FLG* variants and repeated measures of allergic sensitization, on allergic disorders, generalized estimating equations (GEE) were used when estimating the covariance matrix to account for correlated observations and the within-child effect of the repeated measurements [88]. Risk ratios (RR) and their 95% confidence intervals (95% CIs) were estimated by applying log-binomial regression

models using the GENMOD procedure in SAS 9.3 (SAS, Gary, NC, USA). The "no allergic disorder" category formed the reference category. The main effects of *FLG* variants and allergic sensitization on the risk for single allergic disorders and comorbidities were estimated separately (single risk factor models). In a full model, the two main plus their interaction effect (*FLG* variants × allergic sensitization) were estimated. We repeated the aforementioned analysis while excluding the 1-or-2 years follow-up data from the GEE analysis to estimate whether possible misclassifications of phenotypes at early life biased our results. In the case of a possibly statistically significant interaction ($P_{\text{interaction}} < 0.1$), the combined effect was estimated as follows: RR = exp[(β_1 × allergic sensitization) + (β_2 × *FLG* variants) + (β_3 × [allergic sensitization × *FLG* variants])]. The term 'combined effect' was used to describe the joint impact of two individual risk factors plus their interaction on the occurrence of the outcome. In all GEE models, sex and age at follow-up were included as potential confounders. To evaluate statistical significance of main effects of covariates (not interaction terms), a *P* value ≤ 0.05 was used as indicator of statistical significance.

Population attributable fractions (PAFs), defined as the proportion of disease that can be related to a specific risk factor(s) [89], were estimated using the following formula: PAF = p_c × [(RR – 1)/RR], where p_c is the proportion of diseased-individuals that are exposed [89, 90]. This approach yields valid PAF estimates in the presence of confounding and/or effect modification, only when using adjusted RRs [89, 90]. In case of statistically significant interactions between allergic sensitization and *FLG* variants, the RRs for the combined effect were used to estimate PAFs associated with both risk factors. If no interaction was detected, adjusted RRs were used to estimate separate PAFs

for each of the two risk factors and an overall PAF was obtained using the following formula: $PAF_{overall} = 1 - [(1 - PAF_{allergic\ sensitization}) \times (1 - PAF_{\textit{FLG}\ variants})]$ [89].

4.4 Results

A total of 1,377, 1,210, 1,345, and 1,298 children, out of 1,456, had available information on eczema, asthma, and rhinitis diagnosis at ages 1-or-2, 4, 10, and 18 years, respectively (Table 4.1). SPT results were available for 515 (age 1-or-2: only children with any symptoms), 982, 1,036, and 853 children at ages 1-or-2, 4, 10, and 18 years, respectively. The proportion of positive SPTs increased from 20.6% to 41.4% between ages 1-or-2 and 18 years. Although SPT at age 1-or-2 was performed only on symptomatic children, the proportion of sensitized children at ages 1-or-2 and 4 years were similar (20.6% vs. 19.6%). The combined proportion of carriers (heterozygous) of *FLG* variants R501X, 2282del4 and S3247X was 10.3% (118/1150; Table 4.1).

4.4.1 Prevalence of Allergic Disorders

The period prevalence of "eczema only" was between 7.9% and 8.5% during the first 10 years of life and decreased to 5.6% at 18 years of age (Table 4.1). "Asthma only" was most prevalent (10.2%) at 4 years. The period prevalence of "rhinitis only", despite the drop at age 4, demonstrated an increasing trend over time by reaching 20.4% at 18 years of age. The comorbidity of "eczema and asthma" and "eczema and rhinitis" ranged from 0.6% to 2.4% and 0.6% to 3.6%, respectively, between 1-or-2 and 18 years. Having "asthma and rhinitis" was most common (9.0%) at age 18 years. A coexistence of eczema, asthma, and rhinitis was found in approximately 2% over the study period.

4.4.2 Association Analysis: Single Risk Factor Models

FLG variants showed increased risk of "eczema only" (RR = 1.37, $P = 0.101$) and "rhinitis only" (RR = 1.27, $P = 0.064$; Table 4.2) that was not statistically significance, during the first 18 years of life. However, *FLG* variants predisposed to increased risk of having "eczema + asthma" (RR = 2.41, $P < 0.01$), "asthma + rhinitis" (RR = 1.52, $P < 0.01$), and a trend for increased risk of "eczema + rhinitis" (RR = 1.69, $P = 0.077$). Moreover, the risk of "eczema, asthma, and rhinitis" was 3.43 fold-higher among those with *FLG* haploinsufficiency than those with functional *FLG* gene. Accordingly, *FLG* variants showed stronger associations with allergic comorbidities than single disorders.

Allergic sensitization, statistically significantly, increased the risk of "eczema-", "asthma-", and "rhinitis-only" by 1.70-, 2.50-, 2.76-fold over the study period (Table 4.2). In addition, allergic sensitization predisposed to increased risk of having "eczema + asthma" (RR = 6.64, $P < 0.01$), "eczema + rhinitis" (RR = 3.48, $P < 0.01$), and "asthma + rhinitis" (RR = 4.71, $P < 0.01$). In particular, allergic sensitization was associated with an increased risk for the coexistence of "eczema, asthma, and rhinitis" (RR = 9.22, $P < 0.01$).

4.4.3 Association Analysis: Two Risk Factors Models

Simultaneously adjusting for *FLG* variants and allergic sensitization plus their interaction effect allowed us to determine their individual and combined effects. In these adjusted models, *FLG* variants did not show any association with allergic disorders (Table 4.3). In contrast, allergic sensitization was significantly associated with all allergic disorders. *FLG* variants and allergic sensitization showed possible interactive effects for

three of the studied comorbidities. Combined effect of allergic sensitization and *FLG* variants increased the risk of "eczema and asthma" (RR = 13.67, $P < 0.01$), "asthma and rhinitis" (RR = 7.46, $P < 0.01$), "eczema, asthma, and rhinitis" (RR = 23.44, $P < 0.01$; Table 4.3). To estimate whether the possible misclassifications of rhinitis/asthma at age 1-or-2 years influenced our results, we repeated the previous analyses while excluding data from the 1-or-2 years. Comparable results were obtained when we repeated our analysis by only including data from 4, 10, and 18 years follow-ups in the GEE analysis (data not shown).

4.4.4 Sex Differences

To determine whether sex is a potential effect modifier, we ran separate models for boys and girls. The effects of *FLG* variants and allergic sensitization on the risk of having single and multiple morbidities were similar for boys and girls. For instance, *FLG* variants associated with 2.95-fold (95% CI: 1.35 – 6.42, $P < 0.01$) higher risk of having 'eczema + asthma + rhinitis' among boys; whereas, a 3.97-fold (95% CI: 1.96 – 8.05, $P < 0.01$) higher risk was observed among girls ($P_{\text{interaction}}$ for sex = 0.597; detailed data not shown).

4.4.5 Population Attributable Fractions (PAFs)

PAFs related to *FLG* variants and allergic sensitizations were estimated for the different allergic disorders using adjusted RRs from Table 4.3. A larger fraction of allergic disorders in the population can be related to allergic sensitization than to *FLG* variants, when simultaneously adjusting for their effects (Table 4.4). Around 13.7% to

29.5% of single allergic disorders might be related to having both allergic sensitization and *FLG* variants. High PAFs for having "eczema, asthma, and rhinitis" (49%), "eczema + asthma" (33.7%), and "asthma + rhinitis" (21.2) were related to the combined effect of both allergic sensitization and *FLG* variants.

4.5 Discussion

In this well-characterized longitudinal study, allergic comorbidities were prevalent at all ages. For the first time, we assessed the single and combined effects of *FLG* variants and allergic sensitization on the development of single allergic disorders and comorbidities. *FLG* variants, when adjusting for the effect of allergic sensitization, did not show any association with allergic disorders. Against that, allergic sensitization, adjusting for the effect of *FLG* variants, increased the risk of all allergic disorders. Substantial combined effects of both factors were detected for allergic comorbidities. For instance, a 23-fold increased risk for having "eczema, asthma, and rhinitis" (approximately 2% of the cohort) was associated with the presence of both *FLG* variants and allergic sensitization.

Investigating and understanding the role of deviated immune responses, i.e. allergic sensitization, in the atopic diathesis have dominated the research field. However, the discovery of *FLG* variants and their pronounced role in the development of allergic disorders, eczema in particular, has caused a shift in the research paradigm toward understanding the contribution of dysfunctional skin barrier. *FLG* variants are considered to result in a compromised epidermal barrier that is characterized by increased penetration of allergens and microbes and elevated transepidermal water loss [60]. In

eczema, the role of *FLG* variants is well established; however, since *FLG* is not expressed in the respiratory epithelia its contribution to the development of allergic disorders of the airways is not clear [56, 60]. A suggested pathway is that cutaneous sensitization, facilitated by *FLG* variants, could lead to local and systematic inflammation at distant organs (i.e., lung and nasal tissues) [56]. Also the contrary is of importance, since it has been shown that inflammatory Th2 cytokines can also alter the integrity of the skin barrier, even in the absence of genetic variants that disrupt its integrity and function [112]. Hence, the two competing views: "inside-outside" which suggests that immunologic abnormality is the initiating and triggering step in the pathogenesis of allergic disorders and the "outside-inside" which posits that a defective epidermal barrier is the driver with epiphenomenal immunological sequelae, seem to be plausible and to some extent interdependent [77]. Therefore, simultaneously considering the contribution of genetic predisposition toward a defective skin barrier (i.e., *FLG* variants) and the propensity for immune dysregulation (i.e., allergic sensitization) could provide new and improved insights into the etiopathogenesis of allergic disorders.

Previous studies showed that *FLG* variants increased the risk of eczema and predisposed to higher risk of asthma and rhinitis only in the presence of coexisting eczema [59, 66]. Hence, suggesting that *FLG* variants increase the risk of allergic comorbidities, such as 'asthma plus eczema' and 'rhinitis plus eczema'. Our results further support and add to the previously reported findings. However, we showed that *FLG* variants did not pose any increased risk of single and multiple allergic disorders while controlling for the effect of allergic sensitization. But, significant interactive effects of both risk factors were associated with increased risk of allergic comorbidities.

Although previous investigations inconsistently showed that *FLG* variants increased the risk of allergic sensitization [65, 66, 99], we did only observe a weak association between *FLG* variants and allergic sensitization in this cohort [79]. The possibility of an association between *FLG* variants and allergic sensitization cannot be totally excluded; however, since only around 15% of those with allergic sensitization carry *FLG* variants in our cohort, we speculate that any possible indirect effect of *FLG* variants on allergic diseases that goes through allergic sensitization to be minimal. Moreover, the possibility of non-independence between the two factors, also known as gene-environment correlation, might bias the interactive effects [113]. Due to the weak non-significant association between *FLG* variants and allergic sensitization in our cohort, we have little concern that the unlikely non-independence of the two factors might bias the observed interactive effects. In support of our analytical model, we clearly demonstrated in a previous report that *FLG* variants posed no increased risk of eczema in the absences of allergic sensitization, whereas, allergic sensitization increased the risk of eczema independent of *FLG* variants and the combined effect of both factors led to elevated risk above and beyond their individual effects [79]. Similarly, Wang *et al.* showed that the association between *FLG* variants and eczema was limited to a subset of children with IgE levels ≥ 100 kU/L [114]. Hence, we suggest considering the combined effect of both risk factors rather than taking the cause-effect approach could lead to a better risk assessment.

Our estimated PAFs suggested that approximately 21% to 49% of allergic comorbidities in the population may be attributed to the presence of both *FLG* variants and allergic sensitization. In agreement with results of this report, it has been shown that

a large proportion (38%) of allergic comorbidities can be explained by allergic sensitization alone [115]. Although, the contribution of *FLG* variants to the overall PAFs seems to be limited while controlling for the main effect of allergic sensitization, the PAFs associated with the combined effects of both risk factors (i.e., in the presence of possible interactive effects between *FLG* variants and allergic sensitization) is substantial. Our results suggest that *FLG* variants need allergic sensitization as a necessary component to execute its adverse effect. Therefore, early identification of those who are genetically predisposed to a dysfunctional skin barrier and who have increased allergic sensitization propensity may expedite the success of intervention/prevention efforts and measures.

The etiologic approach taken so far by research in the field of *FLG* variants and allergic disorders has substantially improved our understanding of these multifactorial disorders. However, as suggested by Ballardini *et al.*, allergic disorders should not be studied as single entities [109], rather investigating their coexistence could assess in identifying shared underlying risk factors. Hence, we considered all possible and non-overlapping combinations of eczema, asthma, and rhinitis. Through this approach we were able to show that *FLG* variants increased the risk of allergic comorbidities rather than single disorders, suggesting its contribution towards a more severe and complex clinical phenotype. The effect of *FLG* variants when combined with the effect of allergic sensitization predisposed to even higher risk of allergic comorbidities. In contrast to investigating the coexistence of allergic disorders, evidence supporting the progression of events (i.e., allergic march) also exists [116, 117]. For instance, a valid assumption is that eczema predisposes to allergic sensitization, which in turn predicts asthma and rhinitis.

However, evidence supporting both pathways exists: (i) eczema precedes allergic sensitization and (ii) allergic sensitization precedes eczema [118, 119]. In the context of *FLG* variants, we have demonstrated, previously, that the combined effect of *FLG* variants and allergic sensitization predisposed to increased risk of subsequent eczema [79]. Also, we showed that the combination of *FLG* variants and eczema (*FLG* variants alone does not) increased the risk of subsequent allergic sensitization, which was limited to the first 10 years of life. Taken together, these observations suggest that multiple pathways could underlie the development of allergic disorders. Therefore, the development of allergic disorders is more complex and do not necessarily follow the atopic march paradigm. Although our approach does not take into account the sequence of the individual allergic manifestations, results of this report demonstrated that the co-manifestation of allergic disorders is common and *FLG* variants and allergic sensitization are common risk factors.

Major strengths of our study are the repeated phenotyping and the low attrition (ranging from 5% to 17%). *FLG* variants genotypes were available for the majority (80%) of the study population. Previously, we demonstrated that the genotyped sample did not differ from the total cohort with respect to multiple characteristics [79]. Although skin prick tests were performed at the different follow-ups to determine sensitization status, not measuring specific IgE is a limitation to our study. Moreover, since SPTs at ages 1 and 2 years were performed on symptomatic children, we speculate that selection bias leading to underestimated RRs might have been induced. Although proportions of SPT positivity at ages 1-or-2 years (20.6%) and 4 years (19.6%; SPT performed irrespective of symptoms) were similar, RRs associating allergic sensitization at age 1-or-

2 years with outcomes (i.e., eczema, asthma, and rhinitis) at the same age tended to be smaller than RRs relating allergic sensitization at age 4 years with outcomes at the same age (data not shown), since also the reference group had some symptoms. Hence, the estimated RRs at age 1-or-2 might be biased; however, the extent of the possible selection bias is minimal since our results did not noticeably change when we excluded the 1-or-2 years follow-up data from the analyses. Such observations further demonstrate the robustness of the GEE method in providing population averaged estimates (RRs) while accounting for the within-subject correlations of repeated measurements. Misclassification of eczema cases is minimal since a high proportion of subjects showed typical manifestation of eczema in usual locations (antecubital or popliteal fossae, ankles, face or neck for 97% at 1 year, 91% at 2 years, 75% at 4 years, 86% at 10 years and 76% at 18 years) [25]. In addition, the prevalence of eczema in our cohort (13.7% at age 10 years and 12.3% at age 18 years) is comparable to results from ISAAC studies conducted among children aged 13–14 years in the United Kingdom (UK) (14.7% phase one and 10.6% phase three) and other studies conducted in the UK [5, 120, 121].Other, potential limitations are the definition of asthma and rhinitis symptoms in early life. Our asthma conclusion at ages 10 and 18 years followed the ISAAC criteria [4], which was at that time not available for assessments at ages 1, 2, and 4 years. Although slightly different methods were used to define asthma; however, in a previous report we demonstrated that the minor change in asthma definition did not influence the validity of our asthma classification [37]. In addition, since it is difficult to differentiate between infectious and other forms of rhinitis in infancy, the elevated prevalence at 1-or-2 years (6.5% vs. 2.5% at age 4 years) might have been influenced by misclassifying ‘viral induced infectious

rhinitis' [33]. The influence of the possible misclassification of rhinitis at 1-or-2 years on the overall results of the study is minimal since similar results were obtained when we reran our analyses while excluding the 1-or-2 year data.

In conclusion, comorbidities of eczema, asthma, and rhinitis are prevalent and both allergic sensitization and *FLG* variants are common risk factors that predispose to the comorbidity of allergic disorders. The combination of both risk factors was associated with elevated risk of coexisting allergic manifestations. Approximately 21% to 49% of the comorbidities in the population are related to both allergic sensitization and *FLG* variants. The coexistence of allergic disorders might reflect more severe clinical phenotypes that require more and longer medical care. However, investigations of comorbidities and their underlying etiological risk factors are limited. Therefore, additional etiological investigations regarding allergic comorbidities are recommended, as this will allow for a better phenotypic characterization.

4.6 Acknowledgments

We thank participants and their families who have helped us with this project over the last two decades. We would like to acknowledge the help of all the staff at The David Hide Asthma and Allergy Research Centre in undertaking the 18 year and previous assessments of 1989 Isle of Wight birth cohort. The 18 year assessment of the IOW birth cohort and future analyses were funded by grants from the National Institute of Health, USA (R01 HL082925 and R01AI091905).

Table 4.1. Characteristics of the Isle of Wight study population at different ages

	Follow-up at age			
	1-or-2 years	**4 years**	**10 years**	**18 years**
% (n)	**n = 1377**	**n = 1210**	**n = 1345**	**n = 1298**
Sex				
Male	50.8 (700)	51.1 (618)	50.7 (682)	49.6 (644)
No allergic disorders	68.9 (949)	74.1 (897)	61.8 (832)	52.7 (684)
Eczema only	8.5 (117)	7.9 (96)	8.0 (107)	5.6 (73)
Asthma only	4.9 (68)	10.2 (123)	6.3 (84)	5.5 (71)
Rhinitis only	6.5 (89)	2.5 (30)	13.4 (181)	20.4 (265)
Eczema + asthma	1.9 (26)	2.4 (29)	1.0 (14)	0.6 (7)
Eczema + rhinitis	1.8 (25)	0.6 (7)	2.4 (32)	3.6 (47)
Asthma + rhinitis	5.5 (75)	1.3 (16)	4.7 (63)	9.0 (117)
Eczema + asthma + rhinitis	2.0 (28)	1.0 (12)	2.4 (32)	2.6 (34)
Allergic sensitization	**n = 515**	**n = 982**	**n = 1036**	**n = 853**
≥ 1 positive SPT	20.6 (106)	19.6 (192)	26.9 (279)	41.4 (353)
FLG variants*		**n = 1150**		
Yes		10.3 (118)		

SPT: Skin prick test

* *FLG* variants were present in 118 participants of the 1150 who were genotyped. The prevalence of *FLG* variants changes based on the analytical sample.

Table 4.2. Single risk factor models: associations of *FLG* variants and allergic sensitization with allergic conditions: longitudinal analysis from age 1-or-2 to 18 years

	Single risk factor model: *FLG* variants*		**Single risk factor model: Allergic sensitization***	
	***FLG* variants genotype**		**Allergic sensitization**	
	WT (k = 3831)$	**LOF (k = 442)$**	**SPT – (k = 2453)$**	**SPT + (k = 929)$**
No allergic disorders (Ref.), % (k)	64.2 (2459)	54.3 (240)	65.3 (1602)	29.7 (276)
Eczema only, % (k)	7.6 (291)	8.8 (39)	9.5 (234)	9.8 (91)
RR (95% CI)	1.00	1.37 (0.94 – 1.99)	1.00	1.70 (1.36 – 2.12)‡
Asthma only, % (k)	6.6 (253)	6.1 (27)	7.2 (177)	9.7 (90)
RR (95% CI)	1.00	1.07 (0.71 – 1.61)	1.00	2.50 (1.99 – 3.14)‡
Rhinitis only, % (k)	11.2 (428)	12.4 (55)	10.0 (246)	20.8 (193)
RR (95% CI)	1.00	1.27 (0.99 – 1.64)	1.00	2.76 (2.34 – 3.27)‡
Eczema + asthma, % (k)	1.4 (52)	2.9 (13)	1.1 (27)	4.1 (38)
RR (95% CI)	1.00	2.41 (1.28 – 4.57)‡	1.00	6.64 (4.18 – 10.54)‡
Eczema + rhinitis, % (k)	2.0 (76)	2.7 (12)	2.1 (52)	4.2 (39)
RR (95% CI)	1.00	1.69 (0.95 – 3.02)	1.00	3.48 (2.35 – 5.18)‡
Asthma + rhinitis, % (k)	5.4 (205)	6.8 (30)	3.7 (90)	15.1 (140)
RR (95% CI)	1.00	1.52 (1.04 – 2.23)†	1.00	4.72 (3.76 – 5.94)‡
Eczema + asthma + rhinitis, % (k)	1.8 (67)	5.9 (26)	1.0 (25)	6.7 (62)
RR (95% CI)	1.00	3.43 (2.02 – 5.83)‡	1.00	9.22 (5.63 – 15.12)‡

WT: Wild-type; LOF: Loss-of-function; SPT – : Negative skin prick test; SPT +: Positive skin prick test; Ref.: represent the common reference category; k: refers to the total number of repeated measurements.

* In all GEE models sex and age at follow-up were included as potential confounders.

$ The presented numbers in parenthesis (k =) refer to the number of repeated measurements. See Table 4.1 for more information on the number of individuals.

† *P*-value < 0.05

‡ *P*-value < 0.01

Table 4.3. Two risk factors model: associations of *FLG* variants and allergic sensitization with allergic conditions: longitudinal analysis from age 1-or-2 to 18 years

	Two risk factors model: *FLG* variants and allergic sensitization*					
	***FLG* variants genotype**		**Allergic sensitization**		**$RR_{Interaction}$ ($P_{Interaction}$)#**	**Combined effect†**
	WT (k = 2707)$	**LOF (k = 328)$**	**SPT – (k = 2164)$**	**SPT + (k = 871)$**		
No allergic disorders (Ref.), % (k)	56.9 (1540)	46.3 (152)	66.1 (1431)	30.0 (261)		
Eczema only, % (k)	9.3 (252)	10.1 (33)	9.2 (200)	9.8 (85)	0.98	
RR (95% CI)	1.00	1.11 (0.81 – 1.54)	1.00	1.74 (1.38 – 2.19)‡	(0.961)	
Asthma only, % (k)	7.6 (205)	6.4 (21)	6.7 (144)	9.4 (82)	1.05	
RR (95% CI)	1.00	0.99 (0.65 – 1.53)	1.00	2.65 (2.07 – 3.39)‡	(0.926)	
Rhinitis only, % (k)	13.3 (360)	14.6 (48)	10.4 (226)	20.9 (182)	0.85	
RR (95% CI)	1.00	1.21 (0.95 – 1.55)	1.00	2.68 (2.25 – 3.19)‡	(0.538)	
Eczema + asthma, % (k)	1.6 (44)	4.0 (13)	1.0 (22)	4.0 (35)	5.46	
RR (95% CI)	1.00	0.49 (0.08 – 3.03)	1.00	5.12 (2.95 – 8.90)‡	**(0.016)**	13.67 (7.35 – 25.42)‡
Eczema + rhinitis, % (k)	2.5 (68)	3.4 (11)	2.1 (46)	3.8 (33)	1.48	
RR (95% CI)	1.00	1.52 (0.85 – 2.69)	1.00	3.19 (2.09 – 4.89)‡	(0.511)	
Asthma + rhinitis, % (k)	6.7 (182)	8.2 (27)	3.4 (73)	15.6 (136)	2.04	
RR (95% CI)	1.00	0.79 (0.34 – 1.89)	1.00	4.57 (3.52 – 5.93)‡	**(0.070)**	7.46 (5.07 – 10.98)‡
Eczema + asthma + rhinitis, % (k)	2.1 (56)	7.0 (23)	1.0 (22)	6.5 (57)	3.59	
RR (95% CI)	1.00	1.02 (0.25 – 4.16)	1.00	6.43 (3.62 – 11.42)‡	**(0.053)**	23.44 (12.27 – 44.78)‡

WT: Wild-type; LOF: Loss-of-function; SPT – : Negative skin prick test; SPT +: Positive skin prick test; Ref.: represent the common reference category; k: refers to the total number of repeated measurements; $RR_{Interaction}$: Risk ratio associated with the interaction term.

* In all GEE models sex and age at follow-up were included as potential confounders.

P-value < 0.1 was used to indicate presence of possible statistical interaction. If there was no evidence of interaction, the interaction term was removed from the regression model and only the main effects were modeled.

$ The presented numbers in parenthesis (k =) refer to the number of repeated measurements. See Table 4.1 for more information on the number of individuals.

† Combined effect was estimated only in the presence of possibly statistically significant interaction term (*FLG* variants × allergic sensitization; $P_{Interaction} < 0.1$).

‡ *P*-value < 0.01.

Table 4.4. Population attributable fractions of different allergic conditions related to *FLG* variants and allergic sensitization

	***FLG* variants**	**Allergic sensitization**	**Overall**
	PAF (%)[*]	**PAF (%)[*]**	**PAF[†] (%)**
Eczema only	1.2	12.7	13.7
Asthma only	n/a[‡]	22.6	22.6[¶]
Rhinitis only	2.1	28.0	29.5
Eczema + asthma	n/a[$]	n/a[$]	33.7
Eczema + rhinitis	4.8	28.7	32.1
Asthma + rhinitis	n/a[$]	n/a[$]	21.2
Eczema + asthma + rhinitis	n/a[$]	n/a[$]	49.0

PAF: population attributable fraction; n/a: not available

[*] PAFs were estimated using adjusted risk ratios obtained from Table 4.3.

[†] In the presence of statistically significant interaction, overall PAF were estimated using the RR for the combined effect obtained from Table 4.3.

[‡] PAF was not estimated due to RR < 1.00.

[$] PAF was not estimated due to the presence of possible statistical interaction (see Table 4.3).

[¶] This value represents the PAF associated with allergic sensitization only since PAF associated with *FLG* variants was not estimable.

CHAPTER 5

RESULTS III

MANUSCRIPT # 3: DNA METHYLATION OF THE EPIDERMAL DIFFERENTIATION COMPLEX IN RELATION TO ECZEMA STATUS: AN EXPLORATORY STUDY[3]

[3] A. H. Ziyab, W. Karmaus, J. W. Holloway, H. Zhang, S. E. Steck, S. Ewart, S. H. Arshad. To be submitted to *Journal of European Academy of Dermatology and Venereology*.

5.1 ABSTRACT

Background: Genetic variants within genes located across the epidermal differentiation complex (EDC) genomic region have shown associations with eczema risk. However, knowledge on the role of epigenetic mechanisms (e.g., DNA methylation) within the EDC locus on eczema risk is lacking.

Objective: To evaluate associations between DNA methylation of CpG sites across the EDC genomic region with eczema status.

Methods: DNA methylation levels of 256 CpG sites spanning the EDC locus were measured using Illumina's 450K array in three semi-independent cohorts; a subsample of the Isle of Wight birth cohort (n = 367 participants aged 18 years, *discovery cohort*), *replication cohort I* (n = 146; parents of the F2 generation), and *replication cohort II* (n = 94; F2 generation). Associations between eczema status and DNA methylation were first assessed in the discovery cohort using linear regression. Methylation sites showing an association with eczema in the discovery cohort at a false discovery rate p-value < 0.1 were taken forward for replication.

Results: Differential DNA methylation of CpG site cg12048339 (located within promoter of *S100A6* gene) was associated with eczema specifically among female participants of all study cohorts; whereas, aberrant DNA methylation of cg10959711 (located within promoter of *S100A11* gene) associated with eczema among male participants in all study samples.

Conclusion: This exploratory study demonstrated that DNA methylation of the EDC genomic region could be an important factor in the development of eczema in a sex-specific manner. Future studies corroborating our findings are needed.

5.2 INTRODUCTION

The elevated life-time prevalence of eczema among children (up to 30%) and adults (up to 10%) makes eczema a global public-health burden [22]. Immunologic dysregulation and defective epidermal barrier are the main hallmarks of eczema [46]. In last decades, etiologic research have focused on immune-related factors, including genetics of immune system, in the pathogenesis of eczema; however, the discovery, in 2006, of loss-of-function variants in the filaggrin gene (*FLG*) have caused a shift in the research paradigm towards understanding factors that regulate the formation and integrity of the epidermal barrier [53, 68]. Filaggrin haploinsufficiency, defined as partial or complete loss of filaggrin (filament-aggregating protein) protein, is associated with the development of an impaired epidermal barrier that is characterized by increased penetration of environmental substances and elevated transepidermal water loss [93]. As yet, *FLG* variants are the strongest and most replicated genetic risk factors for eczema development [59].

The *FLG* gene is located within a dense cluster of genes encoding structural and regulatory proteins that are essential for keratinocyte differentiation, known as the epidermal differentiation complex (EDC) [76, 81, 82]. The EDC locus, spanning 2 Mb on human chromosomal region 1q21, harbors around 60 genes encoding loricrin and involucrin proteins and gene-family clusters encoding: (i) small proline-rich (SPRR) and late cornified envelope (LCE) proteins, (ii) calcium-binding proteins (S100A), and (iii) S100-fused type proteins (SFTP) [76, 81, 82]. Apart from *FLG* variants, a limited number of studies identified variants in other genes within the EDC that are associated with eczema. For instance, conflicting results as to whether polymorphisms in the hornerin

gene (*HRNR*), located 78 kb from *FLG* gene, are associated with eczema risk have been reported [122, 123]. Regarding *SPRR3*, Marenholz *et al.* showed that a 24-bp deletion in the gene was associated with eczema [124]. However, Stemmler *et al.* did not find evidence of association between eczema and other EDC genetic variants than *FLG* variants [125]. Both, genome-wide linkage and association studies have detected signals of linkage/association between eczema and the EDC genomic locus [122, 126-128]. Morar *et al.* showed that the genetic linkage of eczema to the EDC region (LOD score = 3.57) was not totally diminished after accounting for *FLG* variants (LOD score = 2.03) [129]. Hence, this finding point towards the presence of other genetic variants within the EDC that influence eczema development. On the other hand, the limited success, apart from *FLG* variants, of associating EDC gene variants with eczema is not promising. Moreover, candidate-gene and genome-wide association studies have only explained a small proportion of the high heritability (up to 85%) and clustering of eczema within families [68, 69].

Epigenetic mechanisms (e.g., DNA methylation) may offer a complementary explanation when considering the role of EDC on eczema. DNA methylation addresses heritable genetic regulatory elements that alter gene activity (e.g., expression) without changing the underlying nucleotide sequence of the DNA [130] have been speculated as plausible predisposing factors in the development of eczema and might as well account for the unexplained genetic effects [69]. Currently, there is a lack of knowledge on the role of epigenetic mechanisms, in particular DNA methylation, of the EDC genomic locus in relation to eczema risk. Therefore, a comprehensive DNA methylation profiling of the EDC genomic region could further add understanding for this genomic region in

development of eczema. To this end, in an EDC-wide screening this explorative study aimed to evaluate the association of DNA methylation across the EDC region with eczema status in participants of the Isle of Wight birth cohort (aged 18 years, F1 generation). The analyses further test any significant findings identified in the screening (discovery) in two semi-independent samples (pregnant cohort participant, their partners, and their offspring, the F2 generation).

5.3 Materials and Methods

5.3.1 Participants and Characteristics

The Isle of Wight birth (IOW) cohort, an unselected population-based study, was recruited between January 1989 and February 1990 in the Isle of Wight, UK, to study the natural history and etiology of allergic diseases. After exclusion of adoptions, perinatal deaths, and refusal for follow-up, 1,456 children (95%) were enrolled, with follow-up assessments conducted at 1, 2, 4, 10, and 18 years of age. Ethics approvals were obtained from the Isle of Wight Local Research Ethics Committee at recruitment and for the subsequent follow-ups (06/Q1701/34). Written informed consent was obtained from parents, participants, or both. In this analysis, the *discovery cohort* is based on a subsample (n = 367, aged 18 years) of the IOW cohort that was randomly selected for epigenetic study (Figure 5.1). We consider two cohorts to test the reproducibility of the screening. *Replication cohort I* (n = 146; parents of the F2 generation) includes participants of the IOW cohort (28.8% were part of the discovery cohort) plus new participants. Expecting mothers and their partners were assessed before delivery. *Replication cohort II* (n = 94; F2 generation) includes the F2 newborns of the IOW

cohort participants (Figure 5.1). Infants were followed-up at ages 3-, 6-, and 12-months. In all study samples, eczema was defined as chronic or chronically relapsing, itchy dermatitis lasting more than 6 weeks with characteristic morphology and distribution, following Hanifin and Rajka criteria [84].

5.3.2 DNA METHYLATION PROFILING

For the discovery cohort, DNA was extracted from whole blood collected at age 18 years using a standard salting out procedure [85], and bisulfite-treated for cytosine to thymine conversion using the EZ 96-DNA methylation kit (Zymo Research, CA, USA). Similarly, whole blood of *replication cohort I* participants' and cord blood of *replication cohort II* participants' were collected and DNA was extracted and bisulfite converted following the aforementioned standard procedures. In all samples, Genome-wide DNA methylation was assessed using the Illumina Infinium HumanMethylation450 BeadChip (Illumina, Inc., CA, USA), which interrogates > 484,000 CpG sites associated with approximately 24,000 genes. For the purpose of this report, we restricted our analysis to 256 CpG sites that span the EDC genomic region (location on human chromosome 1: 151958685 – 153628983). Arrays were processed using a standard protocol [86], with multiple identical control samples assigned to each bisulfite conversion batch to assess assay variability and samples randomly distributed on microarrays to control against batch effects [131]. The BeadChips were scanned using a BeadStation, and the methylation level calculated for each queried CpG locus (cytosine-guanine dinucleotides) using the Methylation Module of BeadStudio software. We used the DNA methylation profiles of our participants to predict their cell-type composition via a method proposed

by Houseman *et al.* [132] and validated by Koestler *et al.* [133]. This method identified CpG loci from our data which were within differentially methylated regions (DMRs) known to distinguish between white blood cell types, then utilized our Beta-values to predict the proportions of CD8+ T-cells, CD4+ T-cells, natural killer cells, B-cells, monocytes and granulocytes for each sample.

5.3.3 Statistical Analysis

After performing quality control measures on DNA methylation data, *Beta*-values presented as the proportion of methylated (M) over methylated (M) and unmethylated (U) sites (*Beta*-value = M/[c+M+U], with c being constant to prevent dividing by zero) were calculated for each CpG site [87]. *Beta*-values provide intuitive biological interpretation (i.e., % methylation); however, (i) being constrained between values of 0 (unmethylated) and 1 (completely methylated) and (ii) demonstrating high heteroscedasticity in the lower and upper ends of the methylation range has raised concerns on the validity of this measure when performing statistical analysis [91]. The conversion of *Beta*-values to *M*-values (*M*-value = $\log_2$(*Beta*-value/[1 – *Beta*-value]) has been shown to overcome the limitations of beta-values [91]. Moreover, the use of *M*-values, when performing statistical analysis, demonstrated increased reliability in terms of detecting differentially methylated CpG sites; however, *M*-values cannot be directly interpreted [91, 92]. Therefore, following previous recommendation [91], we used *M*-values when performing the statistical analysis and reported the statistics at the scale of *Beta*-values.

All statistical analyses were conducted using SAS® version 9.3 (SAS Institute, Cary, NC, USA). In the *discovery cohort*, linear regression was used to test the

association between DNA methylation (*M*-values) of all CpG sites spanning the EDC region with eczema status adjusting for sex and cell-type composition. To minimize false-positive findings identified in 256 models, CpG sites with a false discovery rate (FDR) adjusted p-value < 0.1 were selected and considered as potentially associated with eczema. The selected CpG sites, in the *discovery cohort*, were taken forward and tested for their possible association with eczema status in *replication cohorts I and II.* Then, the selected CpG sites were tested for association with eczema in males and females separately.

5.4 RESULTS

The 12-month period prevalence of eczema was 12.6%, 15.6%, and 16.7% in the *discovery cohort*, *replication cohort I*, and *replication cohort II*, respectively (Table 5.1). The period prevalence of eczema was statistically significantly higher among women compared to men in the *discovery cohort* (15.2% vs. 7.4%, p-value = 0.034) and *replication cohort I* (21.1% vs. 5.9%, p-value = 0.017). In contrast, eczema was more common among infant boys (21.1%) in comparison with infant girls (11.8%) in *replication cohort II*.

Epigenome-wide DNA methylation profiling, covering more than 484,000 CpG sites across the genome, was performed in all three study samples. We restricted our analysis to 256 CpG sites that spanned the EDC genomic region. In the *discovery cohort*, screening for possible associations between DNA methylation of CpG sites within the EDC and eczema identified five CpG sites that are differentially methylated between eczema-affected and eczema-free participants with FDR adjusted p-values < 0.1 (Table

5.2). The five identified CpG sites were then tested in women and men participants separately. The sex-specific results showed that DNA methylation of the five CpG sites was statistically significantly associated with eczema in women but not in men (Table 5.2). However, the directions of the effect sizes of three out of the five CpG sites (cg10959711, cg13493250, and cg01910639) were in agreement between women and men. For instance, DNA methylation at cg13493250 locus (located within the *LCE1F* gene promoter region) was, on average, 4% higher in women and by 3% in men, when comparing eczema-affect to eczema-free participants.

The five identified CpG sites were further tested for association with eczema in *replication cohort I* (Table 5.3). None of the CpG sites showed statistically significant association with eczema. The direction of the effect size of two CpG sites (cg10959711 and cg12048339) was in agreement when comparing results of the *discovery cohort* (Table 5.2) with results of *replication cohort I* (Table 5.3). The sex-stratified analyses showed DNA methylation of two CpG sites in females (cg10959711 and cg12048339) and four CpG sites in males (cg10959711, cg13493250, cg03969260, and cg12048339) were in agreement, in regard to effect size direction, between the *discovery cohort* and *replication cohort I*.

Finally, associations between the five CpG sites and eczema were tested in the *replication cohort II* (Table 5.4). In regard to statistical significance, none of the associations were replicated; however, two CpG sites (cg10959711 and cg12048339) showed similar magnitude of association as seen in the *discovery cohort*. In infant girls, DNA methylation of one CpG site (cg12048339) demonstrated similar effect across the *discovery cohort* and *replication cohort II*. In infant boys, DNA methylation of two CpG

sites (cg10959711 and cg01910639) showed similar effects across the *discovery cohort* and *replication cohort II*.

In the analysis that considered both sexes, there were two CpG sites (cg10959711 and cg12048339) that showed agreement in respect to the magnitude of their effect sizes across the three study samples. Furthermore, inspecting the sex-specific analyses showed that one CpG site (cg12048339) among female participants and one CpG site among male participants (cg10959711) demonstrated similar effect sizes across the three study samples. For instance, DNA methylation of cg12048339 was 2% (*discovery cohort* and *replication cohort I*) and 1% (*replication cohort II*) lower in eczema-affected as compared to eczema-free female participants. On the other hand, DNA methylation of cg10959711 locus was 1% (*discovery cohort* and *replication cohort I*) and 3% (*replication cohort II*) higher in eczema-affected as compared to eczema-free male participants.

5.5 Discussion

This explorative study, for the first time, investigated the possible association between DNA methylation across the EDC genomic region with eczema in three semi-independent population based samples. In the *discovery cohort*, DNA methylation of five CpG sites was associated with eczema status. Although results of replicating these associations did not demonstrate statistical significance, two CpG sites (cg10959711 and cg12048339) showed similar effect sizes across the three study samples. Noticeably, DNA methylation of CpG site cg12048339 was associated with eczema specifically among female participants of all study cohorts; whereas, DNA methylation of

cg10959711 associated with eczema among male participants in all study samples. Such results might point towards sex-specific associations between DNA methylation across the EDC locus and eczema. Moreover, DNA methylation within the EDC genomic region might explain some of the genome-wide linkage and association signals that relate the EDC locus with eczema.

Both of the identified sex-specific CpG sites are located within genes that belong to the *S100A* gene-family cluster on the EDC locus. The cg12048339 is located within the promoter region of *S100A6* gene, specifically 1500 base-pairs upstream of transcription start site (TSS1500). Similarly, the cg10959711 is located within the TSS1500 genomic region of *S100A11* gene. The *S100A* genes, overall, share similar gene structure by containing two introns and three exons (exon-2 and exon-3 are translated) and encode proteins that, in general, contain two calcium-binding EF-hand motifs [134, 135]. The S100A proteins are involved in various biological functions and processes, such as cell cycle and differentiation, transcription, motility, and inflammation [134]. The clustering of *S100A* genes within the EDC locus, in addition to evidence supporting their involvement in epidermal differentiation [135], suggests a biologically-plausible role in the pathogenesis of skin-related diseases. For instance, elevated serum levels of S100A8 and S100A9 proteins have been shown to be related to abnormal differentiation of keratinocytes in skin of psoriasis patients [136].

The current state of knowledge suggests that increased methylation of gene promoters is associated with reduced genes expression; whereas, decreased methylation is associated with increased gene expression; however, emerging evidence suggest that such a regulatory role highly depends on the context in which methylation occurs in the

genome and is not limited to methylation of promoter regions [137, 138]. In this study we observed that among female participants with eczema as compared to those without eczema, DNA methylation of cg12048339 (TSS1500 of *S100A11* gene) was reduced, on average, by 1% to 2% across the three study samples. In contrast, DNA methylation of cg10959711 (TSS1500 of *S100A6* gene) among males with eczema as compared to eczema-free participants was increased, on average, by 1% to 3% across the study cohorts. Such results suggest that altered methylation of CpG sites within the EDC locus could serve as markers of eczema in a sex-specific manner.

Existing evidence support the regulatory role of DNA methylation on the activity of *S100A* genes [139]. Although we observed associations between DNA methylation in two CpG sites with eczema, results of this preliminary report should be interpreted with caution. First, DNA methylation, in this report, was not measured in the respective tissue of the disease (i.e., skin); therefore, the extent to which DNA methylation measured in peripheral blood (or cord blood) relate to other tissues and whether can be used as a biomarker for phenotype variation is unclear and an area of current scientific dispute [137, 140]. Second, the discrepancy, apart from the two identified CpG sites, in the results across the three study samples could be due to (i) the differential phenotype expression over age and/or (ii) the dynamic nature of DNA methylation that changes with respect to age and environmental exposures. For instance, previously we observed a sex-reversal in the prevalence of eczema during adolescence period with more girls developing eczema and more boys outgrowing their eczema [25], such a sex-switchover in the risk could be mediated by changes in DNA methylation.

There is no doubt that this study can only serve as a starting point for future studies due to the limited sample sizes of the studied cohorts. The *discovery cohorts* is the largest sample (n = 367) that we studied and from which we obtained the initial significant findings. *Replications cohorts I and II* were smaller in size (n = 146 and n = 94, respectively) and failed to replicate, in terms of statistical significance, results of the *discovery cohort*. The sex-specific analyses were based on even smaller groups. This said, we acknowledge the fact that two out of five CpG sites were in agreement in regard to their effect sizes across the three study cohorts. Moreover, the observed modest effect sizes are not unusual, apart from cancer studies, for studies investigating complex diseases. For example, a 2.07% difference in DNA methylation of the *ORMDL3* gene when comparing asthmatic (87.90%) to non-asthmatic (85.83%) children has been reported [141]. Currently, there is no consensus as to whether modest changes in DNA methylation are as important or less important when compared to larger changes in regard to the development of complex diseases. To this end, previously we demonstrated that an interaction between genetic variants in the *FLG* gene with adjacent DNA methylation predisposed to greater risk of eczema [80]. These results indicated that the association between *FLG* loss-of-function variants and eczema is modulated by DNA methylation. Therefore, simultaneously considering the DNA methylation of the EDC locus and the underlying genetic variants might provide further insights into the genetics/epigenetics of eczema than when considering each factor separately. Taken together, future studies should incorporate DNA methylation, genetic predisposition, gene expression, and proteins levels to provide a better assessment of the risk.

In conclusion, this exploratory study demonstrated that DNA methylation of the EDC genomic region could be an important factor in the development of eczema and might as well explain some of the observed genome-wide linkage and association signals relating the EDC locus with eczema. Interestingly, we observed gender-specific associations between DNA methylation of CpG sites and eczema, which further support the fact that eczema development is different across the two sexes. Future studies confirming these preliminary observations are needed and should obtain DNA methylation profiles from the disease-tissue to increase reliability and relevant of any findings.

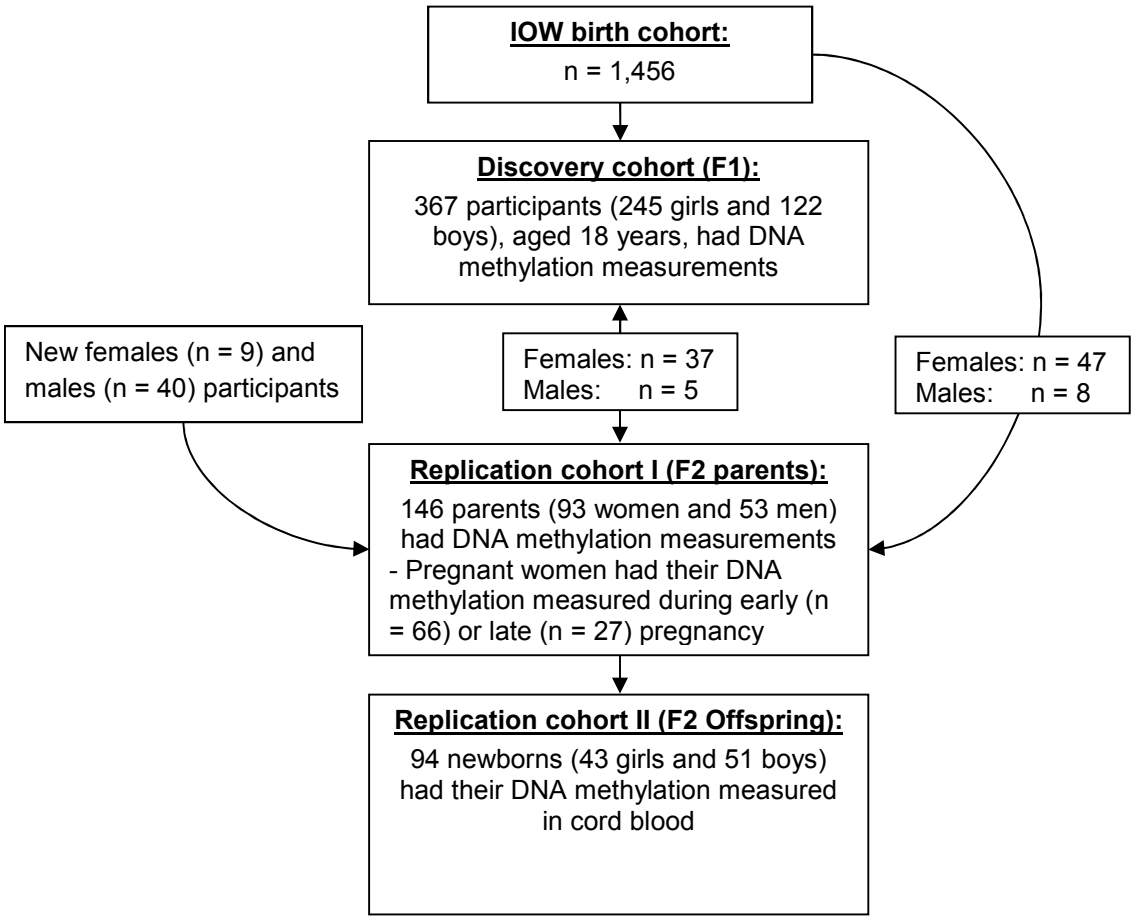

Figure 5.1. Flow diagram of the different analyzed study samples.

Table 5.1. Prevalence of eczema in the different study samples stratified by sex

	Total	Females	Males	p-value*
Discovery cohort (F1)				
Eczema at age 18 years, % (n/total)	12.6 (46/366)	15.2 (37/244)	7.4 (9/122)	0.034
Missing, n	1	1	0	
Replication cohort I (F2 parents)				
Eczema, % (n/total)	15.6 (22/141)	21.1 (19/90)	5.9 (3/51)	0.017
Missing, n	5	3	2	
Replication cohort II (F2 offspring)				
Eczema at age 1 year, % (n/total)	16.7 (12/72)	11.8 (4/34)	21.1 (8/38)	0.354
Missing, n	22	9	13	

* P-value comparing proportions of females to males.

Table 5.2. Association of CpG sites in the EDC genomic region with eczema status: results of the *discovery cohort*

Both sexes				Average methylation (5^{th} – 95^{th} percentile)				
CpG ID	**Gene**	**Placement**	**Average methylation (5^{th} – 95^{th} percentile)**	**Yes eczema**	**No eczema**	**Effect size (SE)***	**Raw p-value**	**FDR adj. p-value**
cg10959711	S100A6	TSS1500	0.11 (0.07 – 0.16)	0.12 (0.09 – 0.19)	0.10 (0.07 – 0.16)	0.02 (0.004)	0.00023	0.055
cg13493250	LCE1F	TSS1500	0.38 (0.29 – 0.49)	0.41 (0.31 – 0.52)	0.38 (0.29 – 0.48)	0.03 (0.010)	0.00052	0.055
cg03969260	SPRR2G	3′ UTR	0.93 (0.90 – 0.95)	0.94 (0.92 – 0.95)	0.93 (0.91 – 0.94)	0.01 (0.002)	0.00065	0.055
cg12048339	S100A11	TSS1500	0.67 (0.62 – 0.72)	0.65 (0.61 – 0.69)	0.67 (0.62 – 0.72)	-0.02 (0.005)	0.00118	0.071
cg01910639	S100A6	Body	0.73 (0.63 – 0.81)	0.75 (0.67 – 0.83)	0.72 (0.62 – 0.80)	0.03 (0.009)	0.00139	0.071
Female participants								
cg10959711	S100A6	TSS1500	0.11 (0.06 – 0.18)	0.12 (0.09 – 0.22)	0.10 (0.06 – 0.17)	0.02 (0.006)	0.0016	
cg13493250	LCE1F	TSS1500	0.38 (0.29 – 0.50)	0.40 (0.31 – 0.52)	0.36 (0.29 – 0.48)	0.04 (0.011)	0.0012	
cg03969260	SPRR2G	3′ UTR	0.93 (0.91 – 0.95)	0.93 (0.92 – 0.95)	0.92 (0.91 – 0.95)	0.01 (0.002)	0.0018	
cg12048339	S100A11	TSS1500	0.67 (0.62 – 0.72)	0.65 (0.60 – 0.70)	0.67 (0.62 – 0.72)	-0.02 (0.005)	0.0002	
cg01910639	S100A6	Body	0.74 (0.62 – 0.82)	0.75 (0.66 – 0.83)	0.73 (0.61 – 0.72)	0.03 (0.012)	0.0066	
Male participants								
cg10959711	S100A6	TSS1500	0.10 (0.08 – 0.13)	0.11 (0.09 – 0.16)	0.10 (0.08 – 0.13)	0.01 (0.005)	0.0251	
cg13493250	LCE1F	TSS1500	0.39 (0.29 – 0.48)	0.42 (0.28 – 0.49)	0.39 (0.29 – 0.48)	0.03 (0.021)	0.1959	
cg03969260	SPRR2G	3′ UTR	0.93 (0.91 – 0.94)	0.92 (0.90 – 0.94)	0.93 (0.91 – 0.94)	-0.01 (0.003)	0.9959	
cg12048339	S100A11	TSS1500	0.66 (0.62 – 0.71)	0.67 (0.64 – 0.69)	0.66 (0.61 – 0.71)	0.01 (0.011)	0.8144	
cg01910639	S100A6	Body	0.71 (0.65 – 0.75)	0.73 (0.67 – 0.76)	0.71 (0.64 – 0.75)	0.02 (0.012)	0.2849	

SE: Standard error; FDR: False discovery rate; Ajd.: adjusted

* Effect size is calculated as average methylation difference between eczema-affected and eczema-free participants using linear regression models. Effect sizes of the 'both sexes' models were adjusted for sex and cell-type composition.

Table 5.3. Association of CpG sites in the EDC genomic region with eczema status: results of the *replication cohort I*

Both sexes				Average methylation (5^{th} – 95^{th} percentile)			
CpG ID	**Gene**	**Placement**	**Average methylation (5^{th} – 95^{th} percentile)**	**Yes eczema**	**No eczema**	**Effect size (SE)***	**Raw p-value**
cg10959711	S100A6	TSS1500	0.09 (0.08 – 0.11)	0.10 (0.08 – 0.11)	0.09 (0.08 – 0.11)	0.01 (0.003)	0.8798
cg13493250	LCE1F	TSS1500	0.39 (0.29 – 0.51)	0.38 (0.29 -0.51)	0.39 (0.31 – 0.51)	-0.01 (0.015)	0.6788
cg03969260	SPRR2G	3′ UTR	0.93 (0.91 – 0.95)	0.92 (0.90 – 0.95)	0.93 (0.91 – 0.95)	-0.01 (0.003)	0.2889
cg12048339	S100A11	TSS1500	0.67 (0.62 – 0.71)	0.65 (0.61 – 0.70)	0.67 (0.62 – 0.71)	-0.02 (0.007)	0.2862
cg01910639	S100A6	Body	0.70 (0.63 – 0.74)	0.69 (0.64 – 0.73)	0.70 (0.63 – 0.74)	-0.01 (0.008)	0.5078
Female participants							
cg10959711	S100A6	TSS1500	0.10 (0.08 – 0.11)	0.10 (0.08 – 0.11)	0.09 (0.08 – 0.11)	0.01 (0.005)	0.7550
cg13493250	LCE1F	TSS1500	0.39 (0.29 – 0.50)	0.37 (0.24 – 0.53)	0.38 (0.29 – 0.50)	-0.01 (0.017)	0.4420
cg03969260	SPRR2G	3′ UTR	0.93 (0.91 – 0.95)	0.92 (0.90 – 0.95)	0.93 (0.91 – 0.95)	-0.01 (0.004)	0.2898
cg12048339	S100A11	TSS1500	0.67 (0.61 – 0.70)	0.65 (0.61 – 0.70)	0.67 (0.62 – 0.70)	-0.02 (0.007)	0.1428
cg01910639	S100A6	Body	0.70 (0.65 – 0.74)	0.69 (0.60 – 0.73)	0.70 (0.66 – 0.75)	-0.01 (0.008)	0.3577
Male participants							
cg10959711	S100A6	TSS1500	0.09 (0.08 – 0.11)	0.10 (0.09 – 0.11)	0.09 (0.08 – 0.11)	0.01 (0.006)	0.7200
cg13493250	LCE1F	TSS1500	0.39 (0.31 – 0.51)	0.40 (0.39 – 0.51)	0.39 (0.31 – 0.51)	0.01 (0.037)	0.4226
cg03969260	SPRR2G	3′ UTR	0.93 (0.91 – 0.94)	0.92 (0.91 – 0.93)	0.93 (0.91 – 0.94)	-0.01 (0.006)	0.9147
cg12048339	S100A11	TSS1500	0.67 (0.62 – 0.72)	0.68 (0.64 – 0.71)	0.67 (0.63 – 0.72)	0.01 (0.019)	0.6263
cg01910639	S100A6	Body	0.68 (0.62 – 0.73)	0.66 (0.65 – 0.72)	0.68 (0.62 – 0.73)	-0.02 (0.023)	0.8495

SE: Standard error; FDR: False discovery rate; Ajd.: adjusted

* Effect size is calculated as average methylation difference between eczema-affected and eczema-free participants using linear regression models. Effect sizes of the 'both sexes' models were adjusted for sex and cell-type composition.

Table 5.4. Association of CpG sites in the EDC genomic region with eczema status: results of the *replication cohort II*

Both sexes				Average methylation (5^{th} – 95^{th} percentile)			
CpG ID	**Gene**	**Placement**	**Average methylation (5^{th} – 95^{th} percentile)**	**Yes eczema**	**No eczema**	**Effect size (SE)***	**Raw p-value**
cg10959711	S100A6	TSS1500	0.19 (0.12 – 0.26)	0.20 (0.09 – 0.33)	0.19 (0.14 – 0.26)	0.01 (0.013)	0.8056
cg13493250	LCE1F	TSS1500	0.38 (0.29 – 0.46)	0.36 (0.25 – 0.45)	0.38 (0.31 – 0.46)	-0.02 (0.015)	0.1453
cg03969260	SPRR2G	3′ UTR	0.92 (0.90 – 0.94)	0.91 (0.90 – 0.95)	0.92 (0.89 – 0.94)	-0.01 (0.004)	0.9833
cg12048339	S100A11	TSS1500	0.68 (0.63 – 0.73)	0.68 (0.63 – 0.73)	0.69 (0.64 – 0.73)	-0.01 (0.009)	0.3941
cg01910639	S100A6	Body	0.85 (0.75 – 0.89)	0.82 (0.53 – 0.90)	0.84 (0.77 – 0.89)	-0.02 (0.017)	0.6029
Female participants							
cg10959711	S100A6	TSS1500	0.19 (0.15 – 0.23)	0.17 (0.08 – 0.21)	0.19 (0.15 – 0.23)	-0.02 (0.015)	0.1489
cg13493250	LCE1F	TSS1500	0.36 (0.27 – 0.45)	0.32 (0.25 – 0.44)	0.36 (0.28 – 0.43)	-0.04 (0.026)	0.1439
cg03969260	SPRR2G	3′ UTR	0.92 (0.89 – 0.94)	0.91 (0.90 – 0.92)	0.92 (0.89 – 0.94)	- 0.01 (0.007)	0.3015
cg12048339	S100A11	TSS1500	0.68 (0.64 – 0.72)	0.67 (0.66 – 0.70)	0.68 (0.64 – 0.72)	-0.01 (0.014)	0.7898
cg01910639	S100A6	Body	0.85 (0.77 – 0.89)	0.83 (0.53 – 0.90)	0.85 (0.80 – 0.89)	-0.02 (0.033)	0.5716
Male participants							
cg10959711	S100A6	TSS1500	0.20 (0.11 – 0.28)	0.24 (0.18 – 0.33)	0.21 (0.11 – 0.28)	0.03 (0.018)	0.1938
cg13493250	LCE1F	TSS1500	0.39 (0.32 – 0.46)	0.39 (0.31 – 0.45)	0.40 (0.32 – 0.48)	-0.01 (0.017)	0.5823
cg03969260	SPRR2G	3′ UTR	0.92 (0.90 – 0.94)	0.92 (0.90 – 0.95)	0.91 (0.90 – 0.93)	0.01 (0.006)	0.5499
cg12048339	S100A11	TSS1500	0.69 (0.63 – 0.74)	0.68 (0.63 – 0.73)	0.70 (0.63 – 0.74)	-0.02 (0.013)	0.4397
cg01910639	S100A6	Body	0.85 (0.75 – 0.89)	0.86 (0.82 – 0.88)	0.84 (0.73 – 0.89)	0.02 (0.016)	0.3421

SE: Standard error; FDR: False discovery rate; Ajd.: adjusted

* Effect size is calculated as average methylation difference between eczema-affected and eczema-free participants using linear regression models. Effect sizes of the 'both sexes' models were adjusted for sex and cell-type composition.

CHAPTER 6

DISCUSSION

6.1 SECTION STRUCTURE

In this section I will tie together the major findings of this dissertation and their relevant interpretation. In general, results of chapters 3 and 4 will by synthesized together as they were (i) based on the same study sample and (ii) investigated the role of *FLG* variants and allergic sensitization on the development of single (chapter 3) and multiple (chapter 4) allergic diseases. In contrast, results of chapter 5 will be discussed separately since it was based on three semi-independent study samples and investigated the potential role of DNA methylation of the EDC genomic locus on the development of eczema.

6.2 OVERVIEW OF FINDINGS

The first two research-related chapters (i.e., chapters 3 and 4) of this dissertation aimed at elucidating the role of *FLG* variants and allergic sensitization on the development of allergic diseases. Chapter 3 focused on the possible effect modification by allergic sensitization and/or eczema on the association between "*FLG* variants and asthma" and "*FLG* variants and rhinitis". I demonstrated that both allergic sensitization and eczema status are possible effect modifiers of the association between *FLG* variants and asthma. In contrast, neither eczema nor allergic sensitization modified the association of *FLG* variants with rhinitis in a statistically significant manner. In chapter 4, I showed that the coexistence of eczema, asthma, and rhinitis is prevalent and allergic sensitization

and *FLG* variants jointly predisposed to increased risk of allergic comorbidities. On the other hand, chapter 5 investigated the role of DNA methylation across the EDC genomic region on the risk of eczema. Results of this exploratory study demonstrated possible gender-specific associations between DNA methylation of CpG sites and eczema in the discovery study sample; however, the results, across the three semi-independent study samples, did not identify consistent pattern of associations between DNA methylation of CpG sites in EDC locus with eczema

6.3 *FLG* Variants and Allergic Sensitization in Relation to Allergic Diseases

Studies investigating the natural history (long-term prognosis) of allergic diseases have improved our understanding of the biological mechanisms that underlie the clinical manifestation of these complex disorders. For instance, in previous reports we have demonstrated that a sex-switchover in the prevalence of eczema [25], asthma [37], and rhinitis [33] occurs around puberty, with more adolescent girls being affect. Etiologic investigations have indicated the importance of environmental, immunologic, and genetic factors in the pathogenesis of allergic diseases [23, 26, 110, 142-144]. Moreover, close and complex inter-relatedness between the manifestations of eczema, asthma, and rhinitis have been demonstrated [45, 48, 103]; however, common mechanisms and pathways underlying their development is an area of ongoing scientific research. The 'allergic march' concept suggests that allergic disorders develop in a sequential pattern, starting with eczema in early infancy and later in childhood developing asthma and rhinitis [41, 106]. However, supporting replicable evidence and consensus is far from complete [108, 145, 146]. An opposing concept is the "coexistence" of allergic morbidities [42, 45, 109,

115]. Therefore, identifying common risk factors that predispose to the development of single and/or multiple allergic diseases is of importance and will further elucidate the etiology of such complex diseases.

Allergic sensitization (deviated immune responses) has been widely associated with different allergic diseases [48, 49, 96]. Similarly, *FLG* variants (representing defective epidermal barrier) have recently emerged as strong independent risk factors for eczema [59]. Furthermore, *FLG* variants increase the risk of asthma and rhinitis only in the presence of coexisting eczema [59, 60]. Apart from eczema, it has been suggested that *FLG* variants predispose to allergic disorders of the airways by facilitating the penetration of allergens through the defective epidermal barrier, which in turn might lead to local and systematic inflammation at distant organs (e.g., lung and nasal tissues) [56, 147]. In this dissertation I have focused on *FLG* variants and allergic sensitization as they have been considered as common risk factors that predispose and link the manifestation of eczema, asthma, and rhinitis [49, 60, 79]. That said, our current state of knowledge lacks clear understanding on their single and joint role in the development of single (chapter 3) and multiple (chapter 4) allergic diseases. Utilizing the prospective nature of the Isle of Wight birth cohort study, I tested if *FLG* variants interact with allergic sensitization and/or eczema on the risk of later asthma and rhinitis. Specifically, I investigated as to whether interplay between "*FLG* variants and allergic sensitization" and/or "*FLG* variants and eczema" predisposed to the development of subsequent asthma and rhinitis. Next, I tested whether *FLG* variants and allergic sensitization jointly predisposed to the coexistence of eczema, asthma, and rhinitis.

6.3.1 Single Allergic Diseases

Thus far, *FLG* variants are the strongest and most replicated genetic risk factors for eczema development [59, 66]. In a previous report, we have corroborated the reported association between *FLG* variants and eczema and further demonstrated that allergic sensitization is an important effect modifier of this association [79]. We showed that *FLG* variants increased the risk of eczema only in the presence of allergic sensitization and the combined effect of both risk factors was associated with elevated risk of eczema (RR = 2.92, 95% CI: 1.47–5.77) that was above and beyond their individual effects [79]. In this dissertation, I sought to explore the possibility that allergic sensitization could also modify the reported associations of *FLG* variants with asthma and rhinitis.

Prior studies have shown that *FLG* variants are associated with asthma only in the presence of coexisting eczema [59, 66]. However, this possible effect modification by eczema was not consistently demonstrated for the association between *FLG* variants and rhinitis [65, 66]. To this end, I have tested the hypothesis that allergic sensitization, rather than eczema status, modifies the association of *FLG* variants with asthma and rhinitis. I have tested my hypothesis under two models: (i) concurrent effect models in which both the exposure and the outcome coexisted at the same time and (ii) delayed effect models in which I used the preceding cumulative allergic sensitization status (or eczema status), i.e. all SPTs results (or eczema status) from assessments that are prior to the outcome assessment, when testing the interaction between *FLG* variants and allergic sensitization (or eczema) on the risk of subsequent asthma/rhinitis.

In stratified analyses (concurrent model), both eczema and allergic sensitization status modulated the association between *FLG* variants and asthma. Statistical

interactions on multiplicative scale were tested to determine whether the observed heterogonous effects were statistically significantly different. A statistically significant interaction was found only between *FLG* variants and allergic sensitization on the risk of asthma. In delayed effect models, which take the time order of risk factors and disease occurrences into account, I have tested the interaction of '*FLG* variants with preceding allergic sensitization' and '*FLG* variants with preceding eczema' on the risk of subsequent asthma. Results of delayed effect models showed that *FLG* variants in interaction with preceding allergic sensitization and preceding eczema increased the risk of subsequent asthma. Hence, indicating that both allergic sensitization and eczema might act as effect modifiers for the association between *FLG* variants and asthma.

In regard to rhinitis, results of stratified analyses (concurrent model) demonstrated that eczema status did not modify the association between *FLG* variants and rhinitis; however, this association was modified when stratifying by allergic sensitization status. But, there was no evidence for an interactive effect (statistical interaction) neither between *FLG* variants and eczema nor between *FLG* variants and allergic sensitization on the risk for rhinitis. Similarly, results of delayed effect models showed that neither preceding eczema nor preceding allergic sensitization interacted with *FLG* variants on the risk of subsequent rhinitis.

These findings indicate that *FLG* variants may need eczema and/or allergic sensitization to execute their adverse effects on asthma. The finding of an interaction between *FLG* variants and allergic sensitization in this study improves our understanding of the possible link between genetics of the epidermal barrier (i.e., *FLG* variants) and a respiratory disorder (i.e., asthma). Such an observation further supports the hypothesis

that cutaneous sensitization priming, partially facilitated by *FLG* variants, could lead to local and systematic inflammation at distant organs (e.g., lung) [56, 98]. Our results of an interactive effect between *FLG* variants and allergic sensitization on the risk of asthma and the lack of such interaction on the risk of rhinitis suggest that the pathway through which *FLG* variants predispose to asthma and rhinitis may be different. For instance, the expression of filaggrin protein was not detected in the human bronchial epithelium; however, it has been shown that filaggrin is expressed in the nasal vestibule [64, 94]. That said, the extent to which filaggrin expression in the nasal vestibular lining influence the development of rhinitis is yet to be investigated by future studies.

In summary, the first part of this dissertation has demonstrated that (i) *FLG* variants are associated with single allergic diseases such as asthma and rhinitis and (ii) allergic sensitization and eczema status act as possible effect modifiers of the association between *FLG* variants and asthma. The combined effects of 'preceding allergic sensitization and *FLG* variants' and 'preceding eczema and *FLG* variants' increased the risk of subsequent asthma by 4.93-fold and 3.33-fold, respectively, during the study period (i.e., the first 18 years of life). In contrast, the association between *FLG* variants and rhinitis was not modulated by either allergic sensitization or eczema status. Taken together, our results demonstrated differential etiological pathways underlying the development of asthma and rhinitis and future studies corroborating our observations are needed.

6.3.2 Multiple Allergic Diseases

After exploring the role of allergic sensitization and *FLG* variants on the development of single allergic diseases, I investigated their single and joint contribution towards the development of allergic comorbidities (i.e., coexistence of eczema, asthma, and rhinitis). First, I constructed eight non-overlapping combinations of the three allergic disorders (Table 2.1), which were used in subsequent analyses for this part of the dissertation.

Although previous studies have demonstrated the importance of allergic sensitization and *FLG* variants on the development of single allergic diseases [49, 60, 96], no prior study have investigated the joint role of both risk factors on the development of allergic comorbidities. In this dissertation, I have demonstrated that the coexistence of "eczema, asthma, and rhinitis" affect around 2% of the study population. Both risk factors, when tested separately, were associated with single and multiple allergic diseases. However, *FLG* variants, when adjusting for the effect of allergic sensitization, did not show any association with allergic disorders. Remarkably, interactions between the two risk factors were associated with substantial combined effects on the risk of allergic comorbidities. For instance, a 23-fold increased risk for having "eczema, asthma, and rhinitis" was associated with the presence of both *FLG* variants and allergic sensitization. Moreover, we estimated that approximately 21% to 49% of allergic comorbidities in the population may be attributed to the presence of both *FLG* variants and allergic sensitization.

In summary, results of this part of the dissertation suggests that comorbidities of eczema, asthma, and rhinitis are prevalent and both allergic sensitization and *FLG*

variants are common risk factors that predispose to the comorbidity of allergic disorders. The combination of both risk factors was associated with elevated risk of coexisting allergic manifestations. Although in the previous section of this dissertation I showed that *FLG* variants and allergic sensitization predispose to single allergic diseases, in this section I further demonstrated the importance of their joint role on the co-manifestation of allergic diseases. Therefore, simultaneously considering the contribution of genetic predisposition toward a defective skin barrier (i.e., *FLG* variants) and the propensity for immune dysregulation (i.e., allergic sensitization) provided new and improved insights into the pathogenesis of allergic disorders.

6.3.3 Strengths and Limitations

The first two research-related chapters (i.e., chapters 3 and 4) in this dissertation were based on information collected in the Isle of Wight birth cohort study. Major strengths of our prospective study, covering the first 18 years of life, are the repeated phenotyping, objective assessments of allergic sensitization status, and the low loss to follow-up (ranged from 5% to 17%). Moreover, the majority (80%) of the study participants were genotyped for *FLG* variants common among populations of European ancestry. We showed, previously, that the genotyped study participants did not differ from the total cohort with regard to multiple characteristics [79]. Hence, there is no indication of selection bias that could pose a threat to the validity of our study. Misclassification of eczema cases is minimal since a high proportion of subjects showed typical manifestation of eczema in the usual locations (antecubital or popliteal fossae,

ankles, face or neck for 97% at 1 year, 91% at 2 years, 75% at 4 years, 86% at 10 years and 76% at 18 years) [25].

Potential limitations are the definition of asthma and rhinitis symptoms in early life. Our asthma conclusion at ages 10 and 18 years followed the ISAAC criteria [4], which was at that time not available for assessments at age 1, 2, and 4 years. Although slightly different methods were used to define asthma, in a previous report we have shown that the minor change in asthma definition over time did not influence the validity of our asthma classification (i.e, no noticeable change in asthma prevalence up to 10 years of age) [37]. In addition, since it is difficult to differentiate between infectious and other forms of rhinitis in infancy, the elevated prevalence at 1-or-2 years might have been influenced by misclassifying 'viral induced infectious rhinitis' [33]. However, by applying repeated measurements analysis based solely on assessments later in childhood and adolescence (4, 10, and 18 years), we believe that the influence of the possible misclassification of rhinitis at 1-or-2 years on the overall results of the study is minimal, since similar results were obtained. Moreover, since SPTs at age 1-or-2 years were performed on symptomatic children, we speculate that selection bias leading to underestimated RRs might have been induced. Although proportions of SPT positivity at ages 1-or-2 years (20.6%) and 4 years (19.6%; SPT performed irrespective of symptoms) were similar, RRs associating allergic sensitization at age 1-or-2 years with outcomes (i.e., eczema, asthma, and rhinitis) at the same age tended to be smaller than RRs relating allergic sensitization at age 4 years with outcomes at the same age, since also the reference group had some symptoms. Hence, the estimated RRs at age 1-or-2 might be underestimated; however, the extent of the possible selection bias is minimal since our

results did not noticeably change when we excluded the 1-or-2 years follow-up data from the analyses. Such observations further demonstrate the robustness of the GEE method in providing population averaged estimates (RRs) while accounting for the within-subject correlations of repeated measurements.

6.4 DNA Methylation of the EDC and Eczema Risk

The third research-related chapter (i.e., chapter 5) of this dissertation focused on the role of DNA methylation across the EDC genomic region in relation to eczema status. In this explorative study, I have screened for associations between DNA methylation of 256 CpG sites across the EDC locus with eczema status at age 18 years (discovery cohort), and subsequently replicated the significant findings in two semi-independent cohorts. Results of this analysis showed that DNA methylation of two CpG sites (cg10959711 and cg12048339) demonstrated similar effect sizes across the three study samples. Noticeably, DNA methylation of CpG site cg12048339 was associated with eczema specifically among female participants of all study cohorts; whereas, DNA methylation of cg10959711 associated with eczema among male participants in all study samples. Such results might point towards sex-specific associations between DNA methylation across the EDC locus and eczema.

Both of the identified sex-specific CpG sites are located within genes that belong to the *S100A* gene-family cluster on the EDC locus. The cg12048339 is located within the promoter region of *S100A6* gene, specifically 1500 base-pairs upstream of transcription start site (TSS1500). Similarly, the cg10959711 is located within the TSS1500 genomic region of *S100A11* gene. The *S100A* genes, overall, share similar gene

structure by containing two introns and three exons (exon-2 and exon-3 are translated) and encode proteins that, in general, contain two calcium-binding EF-hand motifs [134, 135]. The S100A proteins are involved in various biological functions and processes, such as cell cycle and differentiation, transcription, motility, and inflammation [134]. The clustering of *S100A* genes within the EDC locus, in addition to evidence supporting their involvement in epidermal differentiation [135], suggests a biologically-plausible role in the pathogenesis of skin-related diseases. In this study we observed that among female participants with eczema as compared to those without eczema, DNA methylation of cg12048339 (TSS1500 of *S100A11* gene) was reduced, on average, by 1% to 2% across the three study samples. In contrast, DNA methylation of cg10959711 (TSS1500 of *S100A6* gene) among males with eczema as compared to eczema-free participants was increased, on average, by 1% to 3% across the study cohorts. Such results suggest that altered methylation of CpG sites within the EDC locus could serve as markers of eczema in a sex-specific manner.

6.4.1 Strengths and Limitations

There is no doubt that this study can only serve as a starting point for future studies due to the limited sample sizes of the studied cohorts. The *discovery cohorts* is the largest sample (n = 367) that we studied and from which we obtained the initial significant findings. *Replications cohorts I and II* were smaller in size (n = 146 and n = 94, respectively) and failed to replicate, in terms of statistical significance, results of the *discovery cohort*. The sex-specific analyses were based on even smaller groups. Moreover, DNA methylation, in this report, was not measured in the respective tissue of

the disease (i.e., skin); therefore, the extent to which DNA methylation measured in peripheral blood (or cord blood) relate to other tissues and whether can be used as a biomarker for phenotype variation is unclear and an area of current scientific dispute [137, 140]. The discrepancy, apart from the two identified CpG sites, in the results across the three study samples could be due to (i) the differential phenotype expression over age and/or (ii) the dynamic nature of DNA methylation that changes with respect to age and environmental exposures. For instance, previously we observed a sex-reversal in the prevalence of eczema during adolescence period with more girls developing eczema and more boys outgrowing their eczema [25], such a sex-switchover in the risk could be mediated by changes in DNA methylation.

6.5 Conclusion

In this dissertation I have investigated the role of *FLG* variants (representing an impaired epidermal barrier) and allergic sensitizations (representing deviated immune responses) on the development of single and multiple allergic diseases. Results of this dissertation clearly demonstrated that *FLG* variants and allergic sensitization are common factors that predispose and link the manifestation of different allergic diseases (i.e., eczema, asthma, and rhinitis). First, I have shown that allergic sensitization and eczema modulated the association between *FLG* variants and asthma, but not rhinitis. Hence, implying that the mechanisms and pathways through which *FLG* variants predispose to increased risk of asthma and rhinitis may be different. Second, a coexistence of allergic disorders was demonstrated to be frequent; and allergic sensitization and *FLG* variants jointly increased risk of allergic comorbidities. The coexistence of allergic diseases may

represent more severe and complex clinical phenotypes that require more and longer medical care. The interactive effect and the elevated proportion of allergic comorbidities associated with allergic sensitization and *FLG* variants emphasize their joint importance in the pathogenesis of allergic disorders. Third, DNA methylation across the EDC genomic region might influence the risk of eczema in a sex-specific manner.

In summary, this dissertation provided insights on (i) the joint role of *FLG* variants and allergic sensitization on the development of single and multiple allergic diseases and (ii) the role of DNA methylation of the EDC locus on the development of eczema. Therefore, results of this dissertation helped improve our ability to better identify and stratify those who share similar risk characteristics, which may also enhance clinical management and the development of efficient intervention and prevention strategies. Future studies corroborating our findings of the joint role of *FLG* variants and allergic sensitization on the development of single and multiple allergic diseases are needed. Moreover, future studies confirming our preliminary observations of sex-specific associations between DNA methylation and eczema are needed and should obtain DNA methylation profiles from the disease-tissue (i.e., skin) to increase reliability and enhance relevance of any findings.

REFERENCES

1. Pawankar R, Canonica GW, Holgate ST, Lockey RF. Allergic diseases and asthma: a major global health concern. *Curr Opin Allergy Clin Immunol* 2012; **12**:39-41.

2. Suh DC, Sung J, Gause D *et al.* Economic burden of atopic manifestations in patients with atopic dermatitis--analysis of administrative claims. *J Manag Care Pharm* 2007; **13**:778-89.

3. Emerson RM, Williams HC, Allen BR. What is the cost of atopic dermatitis in preschool children? *Br J Dermatol* 2001; **144**:514-22.

4. Asher MI, Keil U, Anderson HR *et al.* International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. *Eur Respir J* 1995; **8**:483-91.

5. Asher MI, Montefort S, Bjorksten B *et al.* Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys. *Lancet* 2006; **368**:733-43.

6. Worldwide variations in the prevalence of asthma symptoms: the International Study of Asthma and Allergies in Childhood (ISAAC). *Eur Respir J* 1998; **12**:315-35.

7. Williams H, Robertson C, Stewart A *et al.* Worldwide variations in the prevalence of symptoms of atopic eczema in the International Study of Asthma and Allergies in Childhood. *J Allergy Clin Immunol* 1999; **103**:125-38.

8. Worldwide variation in prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and atopic eczema: ISAAC. The International Study of Asthma and Allergies in Childhood (ISAAC) Steering Committee. *Lancet* 1998; **351**:1225-32.

9. Pearce N, Ait-Khaled N, Beasley R *et al.* Worldwide trends in the prevalence of asthma symptoms: phase III of the International Study of Asthma and Allergies in Childhood (ISAAC). *Thorax* 2007; **62**:758-66.

10. Asher MI. Recent perspectives on global epidemiology of asthma in childhood. *Allergol Immunopathol (Madr)* 2010; **38**:83-7.

11. Flohr C. Recent perspectives on the global epidemiology of childhood eczema. *Allergol Immunopathol (Madr)* 2011; **39**:174-82.

12. Bousquet J, Khaltaev N, Cruz AA *et al.* Allergic Rhinitis and its Impact on Asthma (ARIA) 2008 update (in collaboration with the World Health Organization, GA(2)LEN and AllerGen). *Allergy* 2008; **63 Suppl 86**:8-160.

13. Deckers IA, McLean S, Linssen S *et al.* Investigating international time trends in the incidence and prevalence of atopic eczema 1990-2010: a systematic review of epidemiological studies. *PLoS One* 2012; **7**:e39803.

14. Odhiambo JA, Williams HC, Clayton TO *et al.* Global variations in prevalence of eczema symptoms in children from ISAAC Phase Three. *J Allergy Clin Immunol* 2009; **124**:1251-8 e23.

15. Anderson HR, Ruggles R, Strachan DP *et al.* Trends in prevalence of symptoms of asthma, hay fever, and eczema in 12-14 year olds in the British Isles, 1995-2002: questionnaire survey. *BMJ* 2004; **328**:1052-3.

16. Lai CK, Beasley R, Crane J *et al.* Global variation in the prevalence and severity of asthma symptoms: phase three of the International Study of Asthma and Allergies in Childhood (ISAAC). *Thorax* 2009; **64**:476-83.

17. Ait-Khaled N, Pearce N, Anderson HR *et al.* Global map of the prevalence of symptoms of rhinoconjunctivitis in children: The International Study of Asthma and Allergies in Childhood (ISAAC) Phase Three. *Allergy* 2009; **64**:123-48.

18. Kauffmann F, Demenais F. Gene-environment interactions in asthma and allergic diseases: challenges and perspectives. *J Allergy Clin Immunol* 2012; **130**:1229-40; quiz 41-2.

19. Bieber T. Atopic dermatitis. *N Engl J Med* 2008; **358**:1483-94.

20. Sohn A, Frankel A, Patel RV, Goldenberg G. Eczema. *Mt Sinai J Med* 2011; **78**:730-9.

21. Sabin BR, Peters N, Peters AT. Chapter 20: Atopic dermatitis. *Allergy Asthma Proc* 2012; **33 Suppl 1**:S67-9.

22. Bieber T. Atopic dermatitis. *Ann Dermatol* 2010; **22**:125-37.

23. Biagini Myers JM, Khurana Hershey GK. Eczema in early life: genetics, the skin barrier, and lessons learned from birth cohort studies. *J Pediatr* 2010; **157**:704-14.

24. Chen W, Mempel M, Schober W, Behrendt H, Ring J. Gender difference, sex hormones, and immediate type hypersensitivity reactions. *Allergy* 2008; **63**:1418-27.

25. Ziyab AH, Raza A, Karmaus W *et al.* Trends in eczema in the first 18 years of life: results from the Isle of Wight 1989 birth cohort study. *Clin Exp Allergy* 2010; **40**:1776-84.

26. Subbarao P, Mandhane PJ, Sears MR. Asthma: epidemiology, etiology and risk factors. *CMAJ* 2009; **181**:E181-90.

27. Herzog R, Cunningham-Rundles S. Pediatric asthma: natural history, assessment, and treatment. *Mt Sinai J Med* 2011; **78**:645-60.

28. Variations in the prevalence of respiratory symptoms, self-reported asthma attacks, and use of asthma medication in the European Community Respiratory Health Survey (ECRHS). *Eur Respir J* 1996; **9**:687-95.

29. Masoli M, Fabian D, Holt S, Beasley R, Global Initiative for Asthma P. The global burden of asthma: executive summary of the GINA Dissemination Committee report. *Allergy* 2004; **59**:469-78.

30. Koterba AP, Saltoun CA. Chapter 9: Asthma classification. *Allergy Asthma Proc* 2012; **33 Suppl 1**:S28-31.

31. Spycher BD, Silverman M, Kuehni CE. Phenotypes of childhood asthma: are they real? *Clin Exp Allergy* 2010; **40**:1130-41.

32. Mucci T, Govindaraj S, Tversky J. Allergic rhinitis. *Mt Sinai J Med* 2011; **78**:634-44.

33. Kurukulaaratchy RJ, Karmaus W, Raza A *et al.* The influence of gender and atopy on the natural history of rhinitis in the first 18 years of life. *Clin Exp Allergy* 2011; **41**:851-9.

34. Kurukulaaratchy RJ, Karmaus W, Arshad SH. Sex and atopy influences on the natural history of rhinitis. *Curr Opin Allergy Clin Immunol* 2012; **12**:7-12.

35. Postma DS. Gender differences in asthma development and progression. *Gend Med* 2007; **4 Suppl B**:S133-46.

36. Vink NM, Postma DS, Schouten JP, Rosmalen JG, Boezen HM. Gender differences in asthma development and remission during transition through puberty: the TRacking Adolescents' Individual Lives Survey (TRAILS) study. *J Allergy Clin Immunol* 2010; **126**:498-504 e1-6.

37. Soto-Ramirez N, Ziyab AH, Karmaus W *et al.* Epidemiologic methods of assessing asthma and wheezing episodes in longitudinal studies: measures of change and stability. *J Epidemiol* 2013; **23**:399-410.

38. Leynaert B, Neukirch F, Demoly P, Bousquet J. Epidemiologic evidence for asthma and rhinitis comorbidity. *J Allergy Clin Immunol* 2000; **106**:S201-5.

39. Bugiani M, Carosso A, Migliore E *et al.* Allergic rhinitis and asthma comorbidity in a survey of young adults in Italy. *Allergy* 2005; **60**:165-70.

40. Demoly P, Bousquet J. The relation between asthma and allergic rhinitis. *Lancet* 2006; **368**:711-3.

41. Spergel JM. From atopic dermatitis to asthma: the atopic march. *Ann Allergy Asthma Immunol* 2010; **105**:99-106; quiz 07-9, 17.

42. Sun HL, Yeh CJ, Ku MS, Lue KH. Coexistence of allergic diseases: patterns and frequencies. *Allergy Asthma Proc* 2012; **33**:e1-4.

43. Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. *Immunol Rev* 2011; **242**:10-30.

44. Yuksel H, Dinc G, Sakar A *et al.* Prevalence and comorbidity of allergic eczema, rhinitis, and asthma in a city in western Turkey. *J Investig Allergol Clin Immunol* 2008; **18**:31-5.

45. Illi S, von Mutius E, Lau S *et al.* The natural course of atopic dermatitis from birth to age 7 years and the association with asthma. *J Allergy Clin Immunol* 2004; **113**:925-31.

46. Boguniewicz M, Leung DY. Atopic dermatitis: a disease of altered skin barrier and immune dysregulation. *Immunol Rev* 2011; **242**:233-46.

47. Torres-Borrego J, Molina-Teran AB, Montes-Mendoza C. Prevalence and associated factors of allergic rhinitis and atopic dermatitis in children. *Allergol Immunopathol (Madr)* 2008; **36**:90-100.

48. Tan RA, Corren J. The relationship of rhinitis and asthma, sinusitis, food allergy, and eczema. *Immunol Allergy Clin North Am* 2011; **31**:481-91.

49. Arshad SH, Tariq SM, Matthews S, Hakim E. Sensitization to common allergens and its association with allergic disorders at age 4 years: a whole population birth cohort study. *Pediatrics* 2001; **108**:E33.

50. Johansson SG, Lundahl J. Asthma, atopy, and IgE: what is the link? *Curr Allergy Asthma Rep* 2001; **1**:89-90.

51. Weidinger S, Baurecht H, Naumann A, Novak N. Genome-wide association studies on IgE regulation: are genetics of IgE also genetics of atopic disease? *Curr Opin Allergy Clin Immunol* 2010; **10**:408-17.

52. Gould HJ, Sutton BJ. IgE in allergy and asthma today. *Nat Rev Immunol* 2008; **8**:205-17.

53. Palmer CN, Irvine AD, Terron-Kwiatkowski A *et al.* Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nat Genet* 2006; **38**:441-6.

54. Kubo A, Nagao K, Amagai M. Epidermal barrier dysfunction and cutaneous sensitization in atopic diseases. *J Clin Invest* 2012; **122**:440-7.

55. Cork MJ, Danby SG, Vasilopoulos Y *et al.* Epidermal barrier dysfunction in atopic dermatitis. *J Invest Dermatol* 2009; **129**:1892-908.

56. O'Regan GM, Irvine AD. The role of filaggrin in the atopic diathesis. *Clin Exp Allergy* 2010; **40**:965-72.

57. Akiyama M. FLG mutations in ichthyosis vulgaris and atopic eczema: spectrum of mutations and population genetics. *Br J Dermatol* 2010; **162**:472-7.

58. Baurecht H, Irvine AD, Novak N *et al.* Toward a major risk factor for atopic eczema: meta-analysis of filaggrin polymorphism data. *J Allergy Clin Immunol* 2007; **120**:1406-12.

59. Rodriguez E, Baurecht H, Herberich E *et al.* Meta-analysis of filaggrin polymorphisms in eczema and asthma: robust risk factors in atopic disease. *J Allergy Clin Immunol* 2009; **123**:1361-70 e7.

60. Irvine AD, McLean WH, Leung DY. Filaggrin mutations associated with skin and allergic diseases. *N Engl J Med* 2011; **365**:1315-27.

61. O'Regan GM, Sandilands A, McLean WH, Irvine AD. Filaggrin in atopic dermatitis. *J Allergy Clin Immunol* 2008; **122**:689-93.

62. Henderson J, Northstone K, Lee SP *et al.* The burden of disease associated with filaggrin mutations: a population-based, longitudinal birth cohort study. *J Allergy Clin Immunol* 2008; **121**:872-7 e9.

63. Marenholz I, Nickel R, Ruschendorf F *et al.* Filaggrin loss-of-function mutations predispose to phenotypes involved in the atopic march. *J Allergy Clin Immunol* 2006; **118**:866-71.

64. Weidinger S, O'Sullivan M, Illig T *et al.* Filaggrin mutations, atopic eczema, hay fever, and asthma in children. *J Allergy Clin Immunol* 2008; **121**:1203-09 e1.

65. Schuttelaar ML, Kerkhof M, Jonkman MF *et al.* Filaggrin mutations in the onset of eczema, sensitization, asthma, hay fever and the interaction with cat exposure. *Allergy* 2009; **64**:1758-65.

66. van den Oord RA, Sheikh A. Filaggrin gene defects and risk of developing allergic sensitisation and allergic disorders: systematic review and meta-analysis. *BMJ* 2009; **339**:b2433.

67. Bussmann C, Weidinger S, Novak N. Genetics of atopic dermatitis. *J Dtsch Dermatol Ges* 2011; **9**:670-6.

68. Brown SJ, McLean WH. Eczema genetics: current state of knowledge and future goals. *J Invest Dermatol* 2009; **129**:543-52.

69. Bataille V, Lens M, Spector TD. The use of the twin model to investigate the genetics and epigenetics of skin diseases with genomic, transcriptomic and methylation data. *J Eur Acad Dermatol Venereol* 2012; **26**:1067-73.

70. van Vliet J, Oates NA, Whitelaw E. Epigenetic mechanisms in the context of complex diseases. *Cell Mol Life Sci* 2007; **64**:1531-8.

71. Talens RP, Boomsma DI, Tobi EW *et al.* Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. *FASEB J* 2010; **24**:3135-44.

72. Heijmans BT, Mill J. Commentary: The seven plagues of epigenetic epidemiology. *Int J Epidemiol* 2012; **41**:74-8.

73. Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. *Cell* 2007; **128**:669-81.

74. Mischke D, Korge BP, Marenholz I, Volz A, Ziegler A. Genes encoding structural proteins of epidermal cornification and S100 calcium-binding proteins form a gene complex ("epidermal differentiation complex") on human chromosome 1q21. *J Invest Dermatol* 1996; **106**:989-92.

75. Segre JA. Epidermal differentiation complex yields a secret: mutations in the cornification protein filaggrin underlie ichthyosis vulgaris. *J Invest Dermatol* 2006; **126**:1202-4.

76. Henry J, Toulza E, Hsu CY *et al.* Update on the epidermal differentiation complex. *Front Biosci (Landmark Ed)* 2012; **17**:1517-32.

77. Wolf R, Wolf D. Abnormal epidermal barrier in the pathogenesis of atopic dermatitis. *Clin Dermatol* 2012; **30**:329-34.

78. Proksch E, Folster-Holst R, Brautigam M *et al.* Role of the epidermal barrier in atopic dermatitis. *J Dtsch Dermatol Ges* 2009; **7**:899-910.

79. Ziyab AH, Karmaus W, Yousefi M *et al.* Interplay of filaggrin loss-of-function variants, allergic sensitization, and eczema in a longitudinal study covering infancy to 18 years of age. *PLoS One* 2012; **7**:e32721.

80. Ziyab AH, Karmaus W, Holloway JW *et al.* DNA methylation of the filaggrin gene adds to the risk of eczema associated with loss-of-function variants. *J Eur Acad Dermatol Venereol* 2013; **27**:e420-3.

81. Kypriotou M, Huber M, Hohl D. The human epidermal differentiation complex: cornified envelope precursors, S100 proteins and the 'fused genes' family. *Exp Dermatol* 2012; **21**:643-9.

82. Hoffjan S, Stemmler S. On the role of the epidermal differentiation complex in ichthyosis vulgaris, atopic dermatitis and psoriasis. *Br J Dermatol* 2007; **157**:441-9.

83. Arshad SH, Karmaus W, Kurukulaaratchy R *et al.* Polymorphisms in the interleukin 13 and GATA binding protein 3 genes and the development of eczema during childhood. *Br J Dermatol* 2008; **158**:1315-22.

84. Hanifin JM, Rajka G. Diagnostic features of atopic dermatitis. *Acta Derm Venereol Suppl (Stockh)* 1980; **92**:44-47.

85. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; **16**:1215.

86. Bibikova M, Barnes B, Tsan C *et al.* High density DNA methylation array with single CpG site resolution. *Genomics* 2011; **98**:288-95.

87. Kuan PF, Wang S, Zhou X, Chu H. A statistical framework for Illumina DNA methylation arrays. *Bioinformatics* 2010; **26**:2849-55.

88. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986; **42**:121-30.

89. Steenland K, Armstrong B. An overview of methods for calculating the burden of disease due to specific risk factors. *Epidemiology* 2006; **17**:512-9.

90. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *Am J Public Health* 1998; **88**:15-9.

91. Du P, Zhang X, Huang CC *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 2010; **11**:587.

92. Bock C. Analysing and interpreting DNA methylation data. *Nat Rev Genet* 2012; **13**:705-19.

93. Brown SJ, McLean WH. One remarkable molecule: filaggrin. *J Invest Dermatol* 2012; **132**:751-62.

94. De Benedetto A, Qualia CM, Baroody FM, Beck LA. Filaggrin expression in oral, nasal, and esophageal mucosa. *J Invest Dermatol* 2008; **128**:1594-7.

95. Ying S, Meng Q, Corrigan CJ, Lee TH. Lack of filaggrin expression in the human bronchial mucosa. *J Allergy Clin Immunol* 2006; **118**:1386-8.

96. Busse WW. The role of allergy in disease. *Immunol Rev* 2011; **242**:5-9.

97. Marenholz I, Kerscher T, Bauerfeind A *et al.* An interaction between filaggrin mutations and early food sensitization improves the prediction of childhood asthma. *J Allergy Clin Immunol* 2009; **123**:911-6.

98. Heimall J, Spergel JM. Filaggrin mutations and atopy: consequences for future therapeutics. *Expert Rev Clin Immunol* 2012; **8**:189-97.

99. Poninska J, Samolinski B, Tomaszewska A *et al.* Filaggrin gene defects are independent risk factors for atopic asthma in a Polish population: a study in ECAP cohort. *PLoS One* 2011; **6**:e16933.

100. Li M, Chen X, Chen R, Bao Y, Yao Z. Filaggrin gene mutations are associated with independent atopic asthma in Chinese patients. *Allergy* 2011; **66**:1616-7.

101. Ekelund E, Lieden A, Link J *et al.* Loss-of-function variants of the filaggrin gene are associated with atopic eczema and associated phenotypes in Swedish families. *Acta Derm Venereol* 2008; **88**:15-9.

102. Greenland S. Interactions in epidemiology: relevance, identification, and estimation. *Epidemiology* 2009; **20**:14-7.

103. Martinez FD, Wright AL, Taussig LM *et al.* Asthma and wheezing in the first six years of life. The Group Health Medical Associates. *N Engl J Med* 1995; **332**:133-8.

104. Reed CE. The natural history of asthma. *J Allergy Clin Immunol* 2006; **118**:543-8; quiz 49-50.

105. Williams HC, Strachan DP. The natural history of childhood eczema: observations from the British 1958 birth cohort study. *Br J Dermatol* 1998; **139**:834-9.

106. Zheng T, Yu J, Oh MH, Zhu Z. The atopic march: progression from atopic dermatitis to allergic rhinitis and asthma. *Allergy Asthma Immunol Res* 2011; **3**:67-73.

107. Barberio G, Pajno GB, Vita D *et al.* Does a 'reverse' atopic march exist? *Allergy* 2008; **63**:1630-32.

108. Williams H, Flohr C. How epidemiology has challenged 3 prevailing concepts about atopic dermatitis. *J Allergy Clin Immunol* 2006; **118**:209-13.

109. Ballardini N, Kull I, Lind T *et al.* Development and comorbidity of eczema, asthma and rhinitis to age 12: data from the BAMSE birth cohort. *Allergy* 2012; **67**:537-44.

110. Holloway JW, Yang IA, Holgate ST. Genetics of allergic disease. *J Allergy Clin Immunol* 2010; **125**:S81-94.

111. Roberts G, Zhang H, Karmaus W *et al.* Trends in cutaneous sensitization in the first 18 years of life: results from the 1989 Isle of Wight birth cohort study. *Clin Exp Allergy* 2012; **42**:1501-9.

112. Gutowska-Owsiak D, Ogg GS. Cytokine regulation of the epidermal barrier. *Clin Exp Allergy* 2013; **43**:586-98.

113. Wagner B, Li J, Liu H, Guo G. Gene-environment correlation: difficulties and a natural experiment-based strategy. *Am J Public Health* 2013; **103 Suppl 1**:S167-73.

114. Wang IJ, Lin TJ, Kuo CF *et al.* Filaggrin polymorphism P478S, IgE level, and atopic phenotypes. *Br J Dermatol* 2011; **164**:791-6.

115. Pinart M, Benet M, Annesi-Maesano I *et al.* Comorbidity of eczema, rhinitis, and asthma in IgE-sensitised and non-IgE-sensitised children in MeDALL: a population-based cohort study. *Lancet Respir Med* 2014; **2**:131-40.

116. van der Hulst AE, Klip H, Brand PL. Risk of developing asthma in young children with atopic eczema: a systematic review. *J Allergy Clin Immunol* 2007; **120**:565-9.

117. Dharmage SC, Lowe AJ, Matheson MC *et al.* Atopic dermatitis and the atopic march revisited. *Allergy* 2014; **69**:17-27.

118. Almqvist C, Li Q, Britton WJ *et al.* Early predictors for developing allergic disease and asthma: examining separate steps in the 'allergic march'. *Clin Exp Allergy* 2007; **37**:1296-302.

119. Lowe AJ, Abramson MJ, Hosking CS *et al.* The temporal sequence of allergic sensitization and onset of infantile eczema. *Clin Exp Allergy* 2007; **37**:536-42.

120. Osman M, Tagiyeva N, Wassall HJ *et al.* Changing trends in sex specific prevalence rates for childhood asthma, eczema, and hay fever. *Pediatr Pulmonol* 2007; **42**:60-5.

121. Burr ML, Dunstan FD, Hand S, Ingram JR, Jones KP. The natural history of eczema from birth to adult life: a cohort study. *Br J Dermatol* 2013; **168**:1339-42.

122. Esparza-Gordillo J, Weidinger S, Folster-Holst R *et al.* A common variant on chromosome 11q13 is associated with atopic dermatitis. *Nat Genet* 2009; **41**:596-601.

123. Greisenegger EK, Zimprich F, Zimprich A, Gleiss A, Kopp T. Association of the chromosome 11q13.5 variant with atopic dermatitis in Austrian patients. *Eur J Dermatol* 2013; **23**:142-5.

124. Marenholz I, Rivera VA, Esparza-Gordillo J *et al.* Association screening in the Epidermal Differentiation Complex (EDC) identifies an SPRR3 repeat number variant as a risk factor for eczema. *J Invest Dermatol* 2011; **131**:1644-9.

125. Stemmler S, Nothnagel M, Parwez Q *et al.* Variation in genes of the epidermal differentiation complex in German atopic dermatitis patients. *Int J Immunogenet* 2009; **36**:217-22.

126. Cookson WO, Ubhi B, Lawrence R *et al.* Genetic linkage of childhood atopic dermatitis to psoriasis susceptibility loci. *Nat Genet* 2001; **27**:372-3.

127. Paternoster L, Zhurov AI, Toma AM *et al.* Genome-wide association study of three-dimensional facial morphology identifies a variant in PAX3 associated with nasion position. *Am J Hum Genet* 2012; **90**:478-85.

128. Weidinger S, Willis-Owen SA, Kamatani Y *et al.* A genome-wide association study of atopic dermatitis identifies loci with overlapping effects on asthma and psoriasis. *Hum Mol Genet* 2013; **22**:4841-56.

129. Morar N, Cookson WO, Harper JI, Moffatt MF. Filaggrin mutations in children with severe atopic dermatitis. *J Invest Dermatol* 2007; **127**:1667-72.

130. Rozek LS, Dolinoy DC, Sartor MA, Omenn GS. Epigenetics: relevance and implications for public health. *Annu Rev Public Health* 2014; **35**:105-22.

131. Soto-Ramirez N, Arshad SH, Holloway JW *et al.* The interaction of genetic variants and DNA methylation of the interleukin-4 receptor gene increase the risk of asthma at age 18 years. *Clin Epigenetics* 2013; **5**:1.

132. Houseman EA, Accomando WP, Koestler DC *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012; **13**:86.

133. Koestler DC, Christensen B, Karagas MR *et al.* Blood-based profiles of DNA methylation predict the underlying distribution of cell types: a validation analysis. *Epigenetics* 2013; **8**:816-26.

134. Marenholz I, Heizmann CW, Fritz G. S100 proteins in mouse and man: from evolution to function and pathology (including an update of the nomenclature). *Biochem Biophys Res Commun* 2004; **322**:1111-22.

135. Eckert RL, Broome AM, Ruse M *et al.* S100 proteins in the epidermis. *J Invest Dermatol* 2004; **123**:23-33.

136. Benoit S, Toksoy A, Ahlmann M *et al.* Elevated serum levels of calcium-binding S100 proteins A8 and A9 reflect disease activity and abnormal differentiation of keratinocytes in psoriasis. *Br J Dermatol* 2006; **155**:62-6.

137. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 2012; **13**:484-92.

138. Shenker N, Flanagan JM. Intragenic DNA methylation: implications of this epigenetic mechanism for cancer research. *Br J Cancer* 2012; **106**:248-53.

139. Lesniak W. Epigenetic regulation of S100 protein expression. *Clin Epigenetics* 2011; **2**:77-83.

140. Terry MB, Delgado-Cruzata L, Vin-Raviv N, Wu HC, Santella RM. DNA methylation in white blood cells: association with risk factors in epidemiologic studies. *Epigenetics* 2011; **6**:828-37.

141. Michel S, Busato F, Genuneit J *et al.* Farm exposure and time trends in early childhood may influence DNA methylation in genes related to asthma and allergy. *Allergy* 2013; **68**:355-64.

142. Flohr C, Mann J. New insights into the epidemiology of childhood atopic dermatitis. *Allergy* 2014; **69**:3-16.

143. Brunekreef B, Von Mutius E, Wong GK *et al.* Early life exposure to farm animals and symptoms of asthma, rhinoconjunctivitis and eczema: an ISAAC Phase Three Study. *Int J Epidemiol* 2012; **41**:753-61.

144. Brunekreef B, Von Mutius E, Wong G *et al.* Exposure to cats and dogs, and symptoms of asthma, rhinoconjunctivitis, and eczema. *Epidemiology* 2012; **23**:742-50.

145. Barberio G, Pajno GB, Vita D *et al.* Does a 'reverse' atopic march exist? *Allergy* 2008; **63**:1630-2.

146. Hopper JL, Bui QM, Erbas B *et al.* Does eczema in infancy cause hay fever, asthma, or both in childhood? Insights from a novel regression model of sibling data. *J Allergy Clin Immunol* 2012; **130**:1117-22 e1.

147. De Benedetto A, Kubo A, Beck LA. Skin barrier disruption: a requirement for allergen sensitization? *J Invest Dermatol* 2012; **132**:949-63.

APPENDIX A – PERMISSION TO REPRINT

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Jun 01, 2014

This is a License Agreement between Ali H Ziyab ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3400311068211
License date	Jun 01, 2014
Licensed content publisher	John Wiley and Sons
Licensed content publication	Clinical & Experimental Allergy
Licensed content title	Allergic sensitization and filaggrin variants predispose to the comorbidity of eczema, asthma, and rhinitis: results from the Isle of Wight birth cohort
Licensed copyright line	This article is protected by copyright. All rights reserved.
Licensed content author	Ali H. Ziyab,Wilfried Karmaus,Hongmei Zhang,John W. Holloway,Susan E. Steck,Susan Ewart,Syed Hasan Arshad
Licensed content date	Apr 8, 2014
Start page	n/a
End page	n/a
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No

Title of your thesis / dissertation	Role of Allergic Sensitization, Filaggrin Variants, and DNA Methylation on the Risk of Allergic Disorders
Expected completion date	Aug 2014
Expected size (number of pages)	125
Total	0.00 USD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a"Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking ‘accept’ in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your Rightslink account (these are available at any time at http://myaccount.copyright.com).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.
- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this licence must be completed within two years of the date of the grant of this licence (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.
- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner. You may not alter, remove or suppress in any manner any copyright,

trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto.

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU

- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.
- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of

Creative Commons Licenses:: Creative Commons Attribution (CC-BY) license Creative Commons Attribution Non-Commercial (CC-BY-NC) license and Creative Commons Attribution Non-Commercial-NoDerivs (CC-BY-NC-ND) License. The license type is clearly identified on the article.

Copyright in any research article in a journal published as Open Access under a Creative Commons License is retained by the author(s). Authors grant Wiley a license to publish the article and identify itself as the original publisher. Authors also grant any third party the right to use the article freely as long as its integrity is maintained and its original authors, citation details and publisher are identified as follows: [Title of Article/Author/Journal Title and Volume/Issue. Copyright (c) [year] [copyright owner as specified in the Journal]. Links to the final article on Wiley's website are encouraged where applicable.

The Creative Commons Attribution License

The Creative Commons Attribution License (CC-BY) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-commercial re-use of an open access article, as long as the author is properly attributed.

The Creative Commons Attribution License does not affect the moral rights of authors, including without limitation the right not to have their work subjected to derogatory treatment. It also does not affect any other rights held by authors or third parties in the article, including without limitation the rights of privacy and publicity. Use of the article must not assert or imply, whether implicitly or explicitly, any connection with, endorsement or sponsorship of such use by the author, publisher or any other party associated with the article.

For any reuse or distribution, users must include the copyright notice and make clear to others that the article is made available under a Creative Commons Attribution license, linking to the relevant Creative Commons web page.

To the fullest extent permitted by applicable law, the article is made available as is and without representation or warranties of any kind whether express, implied, statutory or otherwise and including, without limitation, warranties of title, merchantability, fitness for a particular purpose, non-infringement, absence of defects, accuracy, or the presence or absence of errors.

Creative Commons Attribution Non-Commercial License

The Creative Commons Attribution Non-Commercial (CC-BY-NC) License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

- Use of article content (other than normal quotations with appropriate citation) by for-profit organisations for promotional purposes
- Linking to article content in e-mails redistributed for promotional, marketing or educational purposes;
- Use for the purposes of monetary reward by means of sale, resale, licence, loan, transfer or other form of commercial exploitation such as marketing products
- Print reprints of Wiley Open Access articles can be purchased from:corporatesales@wiley.com

Further details can be found on Wiley Online Libraryhttp://olabout.wiley.com/WileyCDA/Section/id-410895.html

Other Terms and Conditions:

v1.9

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number 501317106. Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support:customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.