EPIDEMIOLOGY AND GENETICS OF ATOPIC DERMATITIS IN SINGAPORE

BANI KAUR SURI

NATIONAL UNIVERSITY OF SINGAPORE

2014

EPIDEMIOLOGY AND GENETICS OF ATOPIC DERMATITIS IN SINGAPORE

BANI KAUR SURI

M. Sc., University of Mumbai, India

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOLOGICAL SCIENCES

NATIONAL UNIVERSITY OF SINGAPORE

2014

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

Bani Kaur Suri

23rd Jan 2014

Acknowledgements

I would like to express my deepest gratitude towards my supervisor A/P Chew Fook Tim, without whom this project would have been impossible. I am grateful for his guidance, support and encouragement throughout the course of my graduate studies.

I wish to thank my former and current friends from lab - Ramani, Pallavi, Tan Ching, Ken, Louis, Jiang Nan, Sravanthy, Anand, Anusha, Yang Yie, Wei Yi and Mdm Oh who in their own ways made the lab a fun-place to work in. I'd like to take this opportunity to thank Nakul and Jeremy for taking time out from their busy schedules to help me.

I am grateful to my wonderful friends- Juzar, Amna and Whiskoo, who formed pillars of support throughout my PhD journey. Friends who were family away from home, without whom this journey seemed impossible!

I would like to thank the fun-loving bunch of friends I have known since my first few months in Singapore, a list too big to name here. In addition, I would also like to thank the badminton group for the wonderful time spent. And a big thank you to all my friends who volunteered every time I needed samples for optimisation.

Finally, I wish to thank and dedicate this thesis to my family for their constant love, encouragement and support. Thank you Mom, Dad, Sunny, Nainee, Sethis and most importantly Darji.

Table of Contents

Summary

Atopic dermatitis (AD), commonly known as eczema, is a chronic itchy skin condition with an increase incidence of occurrence at skin folds. It manifests due to a complex interplay of genetic and environmental factors. The increasing prevalence in industrialised countries such as Singapore has made AD an important health problem worth researching. With this in view, we embarked on a journey; (i) to conduct an epidemiological study for identifying and evaluating risk factors associated with eczema in Singapore, (ii) to screen and validate associations of epidermal barrier-related genes to AD in ethnic Chinese of Singapore, (iii) to functionally characterise *CA6*, a gene found associated with the maximum number of tags and (iv) to replicate previously reported associations in ethnic Chinese.

Our epidemiological study compared to the previously conducted two studies, showed an increase in prevalence of chronic rash at skin flexures to 15.7% compared to 10% in 2001(Wang *et al.*, 2004) and 6.5% in 1994 (Goh *et al.*, 1996) . There has also been an overall increase in self reported eczema. The steep increase in prevalence is indicative of non-genetic (environmental) contributing factors its pathogenesis in Singapore. This study identified demographic variables such as gender, race and age to be associated with eczema, with females, Malays and older age group reporting a higher prevalence. Paternal eczema was found to be the most highly associated [adjusted OR = 3.68 (2.91-4.67, 95% CI), $P< 0.001$], followed by eczema in both parents [adjusted OR = 3.31 (1.91-5..767, 95% CI), $P< 0.001$] confirming

the strong genetic link. Atopy, diet choices and sedentary lifestyle were also found associated after controlling for the demographic confounders.

The strongest association to parental history justified a genetic study based on which a case-control genetic association study was initiated. Candidate genes selected on the basis of the barrier dysfunction hypothesis went through a two-stage analyses using Illumina's BeadXpress platform, which validated association of 27 genes. The gene with the maximum number of representative SNPs associated with AD i.e., carbonic anhydrase VI (*CA6)* was selected for further functional characterisation.

Gene expression and selective exon expression were evaluated for the associated SNPs of *CA6* using Affymetrix's Human Exon ST1.0 array. To evaluated the regulatory effect of methylation, we used Illumina's Infinium HumanMethylation450 BeadChip. SNPs from *CA6* were found associated with gene expression and selective exon expressions indicating alternative splicing. The SNPs were also found to be associated with methylation-regulated gene expression and alternative splicing. These *ex vivo* results prompted us to confirm the findings *in vitro*, experiments for which are under-way. Association was also found to skin pH when compared to one of the significant SNPs suggesting the function of CA6 in pH maintenance. As explained in great details in the thesis, pH regulates the fine balance of skin shedding (desquamation), an aberration in which can compromise the integrity of the skin making an atopic person susceptible to AD. As maintenance of pH is crucial to the maintenance of homeostasis in the epidermis, *CA6* may play a crucial role in the manifestation of AD. This necessitates further research for

understanding the less-known *CA6* in the context of skin and further characterisation of the changes caused by the associated variants.

Simultaneously, we also undertook replication studies for well known associations reported in ethnic Chinese. The SNPs found associated through a genome-wide association study in Han Chinese were replicated in our Singapore cohort, resulting in a paper. Also, the recently found diverse landscape of Filaggrin (*FLG)* mutations in Singapore Chinese was tested for association to AD. We found the common variants of *FLG* associated with the more severe and chronic forms of AD in our cohort.

Overall, this thesis identifies important epidemiological and genetic factors associated with AD. The findings of this thesis pave way for further research of the associated genes, especially highlighting the novel *CA6,* which maybe the silver bullet for AD. Although still in the initial stages of evaluating *CA6's* contribution to AD, this thesis has laid the ground-work for more focussed basic, diagnostic and therapeutic research.

List of Tables

List of Figures

Abbreviations

AD- Atopic dermatitis

AMP- Antimicrobial peptides

ANOVA- Analysis of variance

APC- Antigen presenting cells

AR- Allergic rhinitis

CA6- Carbonic anhydrase VI

CEU- Utah residents with ancestry from northern and western Europe

CHB- Han Chinese in Beijing

Chr- Chromosome

CI- Confidence interval

CNV- Copy-number variant

DMEM- Dulbecco's Modified Eagle Medium

E- Exon

FBS- Foetal bovine serum

FFA- Free fatty acids

FLG- Filaggrin

GWAS- Genome wide association

HapMap- Haplotype Map

IgE- Immunoglobulin E

IL- Interleukin

IMB- Institute of Medical Biology

IRB- Institutional Review Board

ISAAC- International Study of Asthma and Allergies in Childhood

IV- Ichthyosis vulgaris

JPT- Japanese in Tokyo

KKH- KK Women's and Children's Hospital

KLK- Kallikrein

LB- Lamellar body

LD- Linkage disequilibrium

MAF- Minor allele frequency

NMF- Natural moisturizing factor

Not applicable- NA

Not significant- NS

NSC- National Skin Centre

NUS- National University of Singapore

OR- Odds ratio

RE- Restriction enzyme

Ref- Reference

K- S$1000

SC- Stratum corneum

SCORAD- SCORing Atopic Dermatitis

SG- Stratum granulosum

SGVP- Singapore Genome Variation Project

sig- Significant

SNP- Single-nucleotide polymorphism

SPSS- Statistical package for social sciences

SPT- Skin prick test

TEWL- Trans-epidermal water loss

t-SNP- Tag- SNP

UTR- Untranslated region

Vc- Vehicle control

VNTR- Variable-number tandem repeats

WAO- World Allergy Organisation

YRI- Yoruba in Ibadan, Nigeria

Chapter 1

Introduction to Atopic Dermatitis

"I have approximate answers and possible beliefs,
in different degrees of certainty, about different things.
But I'm not absolutely sure of anything."
-Richard Phillips Feynman

1.1 Atopic Dermatitis

Atopic dermatitis (AD) is one of the three classical allergic diseases (asthma, allergic rhinitis and AD) widely studied. AD commonly known as eczema is a relapsing inflammatory skin condition with a predisposition for skin flexures (Williams, 2005). Typical symptoms include xerosis (dryness), pruritus (itch), and erythematous lesions with increased trans-epidermal water loss (TEWL) (Cork *et al.*, 2009). The 'World Allergy Organisation' (WAO) in 2003 revised allergy-related nomenclature and termed eczema as an umbrella term for dermatitis with certain common clinical characteristics involving genetically determined skin barrier defect. It also added that only if the affected person is atopic—which may be concluded by measuring immunoglobulin E (IgE) levels or performing a skin prick test, only then the disease may be termed as atopic eczema (Johansson *et al.*, 2004) or atopic dermatitis. Although not a debilitating disease, AD still has a considerable impact on the patients and their families in terms of quality-of-life and economic burden (Akdis *et al.*, 2006; Delea *et al.*, 2007)

1.2 Clinical phenotype

Clinical presentation of AD differs with age and the course of the disease. Eczematous lesions may be present in acute (oozing, crusted vesicles, erythematous plaques), subacute (thick and excoriated plaques) or chronic (lichenified, pigmented, excoriated plaques) forms. The clinical features of AD during infancy are papulo-vesicular eczematous lesions, mainly present on the cheeks, scalp and other exposed areas. During childhood the flexural areas (ante-cubital fossae, neck wrist, etc), nape of the neck, dorsum of feet and hands get affected while in the adolescent and adult stages lichenified plaques may affect the flexures and neck areas with itch being associated at all 3 stages (Bieber, 2010).

1.3 Prevalence

AD is an early onset disease with over 70% of cases starting under 5 years of age and around 10% reported to start in adulthood (Williams, 2005). Studies have reported childhood prevalence of AD in Europe to be around 10-15% with a persistence rate of 40-60% post puberty (Wüthrich, 2006). It has a high prevalence rate in developed countries like UK (19.5%), Denmark (22.9%), Hong Kong (20.1%), and Japan (19%) (Tay *et al.*, 2002). According to the International Study of Asthma and Allergies in Childhood (ISAAC), the prevalence varies from less than 2 percent in Iran and China to approximately 20 percent in Australia, Scandinavia and United States (Williams, 2005). Surveys conducted in over 30 schools in Singapore in 1994 showed the prevalence of AD with flexural involvement at 6.9% (Goh *et al.*, 1996) which grew to 10.2% in 2001. Amongst the three allergic diseases evaluated, AD was the only one to show a significant increase within 7 years (Wang *et al.*,

2004). In a different cross-sectional epidemiological study, the one year prevalence of eczema in Singapore was reported at 20.8% (Tay *et al.*, 2002). **The reported steep increase in prevalence and severity of AD in the Singapore population motivated us to comprehensively study and analyse this disease in the Singapore context.**

1.4 Causes

AD like many other allergic diseases has a genetic predisposition. Population studies done in Avon county, England showed a strong association between parental eczema and childhood AD with an odds ratio of 2.72 (95% CI: 2.09 to 3.53) when both parents are affected and odds of 1.7 (95% CI: 1.4 to 2.0) with 1 affected parent (Wadonda-Kabondo, 2004). Twin studies reporting concordance rates of 0.72 - 0.86 in monozygotic and 0.21 - 0.23 in dizygotic twins further confirmed the genetic contribution to the disease (Schultz Larsen, 1993; Thomsen *et al.*, 2007). However, genetics alone cannot explain the results of migrant-population studies which show Jamaican children living in London are twice as likely to develop AD as Jamaican children living in Jamaica or the increased risk of AD in smaller families and higher social classes or its rising prevalence in some countries (Williams, 2005). Epidemiological data from Northern Europe has shown a 5-10 fold increase in prevalence rate over four decades reaching a staggering 20% (Schultz Larsen *et al.*, 1996; Schultz Larsen and Hanifin, 1992). Similarly, the prevalence of AD in Swedish schoolchildren has more than doubled between 1979 and 1991 (Broberg *et al.*, 2000). Similar trend observed in Singapore elucidates the importance of environmental factors such as allergens e.g. house-dust mites, food, etc in mediating AD expression. Non-allergic factors such as rough

clothing, *Staphylococcus aureus* infections, exposure to microbes during infancy, excessive heat, and exposure to irritants that disrupt the function of the skin barrier may also be important in disease causation (Williams, 2005). Also, adding to the complexity are ethnic variations. Epidemiological surveys conducted in Singapore and Malaysia showed a higher prevalence of AD in Chinese and Malays as compared to Indians (Tay *et al.*, 2002). This is attributed to the ethnic differences, due to which one interacts differently with a given environment leading to differences in disease expression. These and many more studies indicate **AD as a highly complex disease which involves a complex interplay of genetic and environmental factors in disease manifestation**.

1.5 Different pathogenic models

Atopic dermatitis as the name suggests, has largely been considered an atopic disorder. **Atopy** as defined by WAO is a personal and/or familial tendency, usually in childhood or adolescence, to become sensitized and produce IgE antibodies in response to ordinary exposures to allergens, usually proteins (Johansson *et al.*, 2004). Also, diagnosis of AD in early childhood is considered the start of an 'atopic march' as most of these children later develop asthma and/or other allergic disorders (Kraft *et al.*, 1998). This suggests that AD manifests because of an immunological dysfunction which entails a dominant Th2 pathway, brief mechanism of which is explained below.

Manifestation of type-1 allergic hypersensitivity reaction is a complex pathway involving the Th2 pathway and IgE. The primary exposure known as

sensitisation occurs when an innocuous allergen enters the skin and is internalised by antigen presenting cells (APCs), which are present in the skin. After digestion and processing into its constitutive peptides, the antigen are presented via major histocompatibility complex class II on APC's cell surface. These APCs then migrate and present the antigen to the naive T cells which leads to the maturation of the T cell. The milieu in which the T cells mature shifts the balance to either Th1 or Th2 pathway, which depends on the type of APC presenting the antigen and the genetic predisposition of an individual. An environment of interleukin (IL) 4 (IL4) and IL10 tips the balance to Th2. Following this polarisation, the mature T cells then secrete the various Th2 cytokines including IL3, IL4, IL5, IL9, IL13. This cocktail of cytokines and co-stimulatory molecules recruits and class switches B cells to start producing IgE antibodies. These IgE molecules circulate and bind to high affinity receptors on the resident mast cells of the skin epithelium. Upon the subsequent exposure of the same allergen, the antigens are recognised by the IgE armed mast cells leading to cross-linking and de-granulation of the mast cells spilling the contents like histamine, leukotrienes, super-oxides amongst others leading to inflammatory reaction and tissue damage (Galli *et al.*, 2008). The above described atopic reaction on the skin with elevated IgE levels leading to skin cell disruption as a fallout of inflammation was long thought to be *the* causative mechanism for AD.

A possible autoimmune role of IgE targeting epithelia in AD has also been suggested (Valenta *et al.,* 1998). **Auto-allergy** can be mediated through either (i) cross-linking of cell bound IgE antibodies by auto-allergens leading to release of inflammatory mediators, or (ii) IgE mediated auto-allergen

presentation activating auto-reactive T cells to release pro-inflammatory cytokines, contributing to lesions forming AD (Valenta and R., 2000)**.**

Also, different studies indicate that immune-hyper-reactivity isn't present in all AD patients with over 50% of cases not showing elevated levels of IgE up to age 2 (Illi *et al.*, 2004). However, eventually 80% of these non-atopic patients **(intrinsic AD)** eventually turn atopic (**extrinsic AD**) indicating a non-immune aetiology as the primary event in AD (Cork *et al.*, 2009). Not all children with AD develop the other diseases in the atopic march suggests that elevated IgE levels alone may not be responsible for AD (Kraft *et al.*, 1998). Thus, IgE, a marker for T-cell dysregulation towards the production of Th_2 cytokines, is at best associated with atopic status and not causative of lesions on target organs like skin (Taieb, 1999). Validation of this claim was further strengthened by a meta-analysis study showing a strong association of atopy to disease severity rather than to just AD (Flohr *et al.*, 2004).

Many studies have shown AD severity paralleling permeability barrier abnormalities (Angelova-Fischer *et al.*, 2005; Seidenari and Giusti, 1995; Sugarman *et al.*, 2003). Also different groups have shown both clinically uninvolved skin sites and skin cleared of inflammation for as long as 5 years continue to display significant barrier abnormalities (Eberlein-Konig *et al.*, 2000; Seidenari and Giusti, 1995). Barrier maintenance and restoration through emollient therapy and lipid replenishing creams have shown corrective measures against barrier abnormality along with site inflammation for AD (Cork *et al.*, 2003; Elias *et al.*, 2008).

The gaps identified with the atopic pathogenic model in the early 1990s led Hideoki Ogawa to propose a paradigm shifting hypothesis in his 15th Annual Meeting of Japanese Society for Investigative Dermatology address. He was the very first to hypothesize the importance of a **dysfunctional epidermal barrier** as the primary event in the pathogenesis of AD. He along with Takashi Yoshiike was the first to publish this alternative hypothesis integrating barrier dysfunction and atopy as key players to the pathogenesis of AD (Ogawa and Yoshiike, 1993). Following this, Alain Taeib and others started describing AD in this new perspective which lead to a wider acceptance of this alternative theory (Taieb, 1999). Reports showing reduced levels of — ceramides that hold the cells together (Imokawa *et al.*, 1991), filaggrin which helps retain moisture (Seguchi *et al.*, 1996), transepidermal water loss which indicates penetrability of the skin (Loden, 1995) further corroborated the innate defect of the epidermal skin barrier as the primary event rather than an epiphenomenon in the pathogenesis of AD.

1.6 Skin as a barrier

Skin is the largest organ separating us from the environment. It forms the first line of innate defence by providing not just a physical but also a complex chemical and biological barrier. These defensive functions include: the permeability barrier which retards trans-cutaneous evaporative water loss—allowing survival in a desiccating external environment; antimicrobial barrier which resists growth of microbial pathogens by encouraging colonization by non-pathogenic 'normal' flora (Elias and Schmuth, 2009). Thus, an intact-uncompromised epidermis is a prerequisite for the skin to function as a barrier.

Maintaining the integrity of the epidermis is a very complex process which involves a well regulated squamation (new cell growth) and desquamation (shedding) process for maintaining homeostasis. A flaw in any of the mechanisms maintaining the integrity can lead to a compromised epidermis. This compromised state can be an entry point for allergens and pathogens which can elicit an immune response leading to AD. To understand the barrier we first need to look at the skin and understand its maintenance.

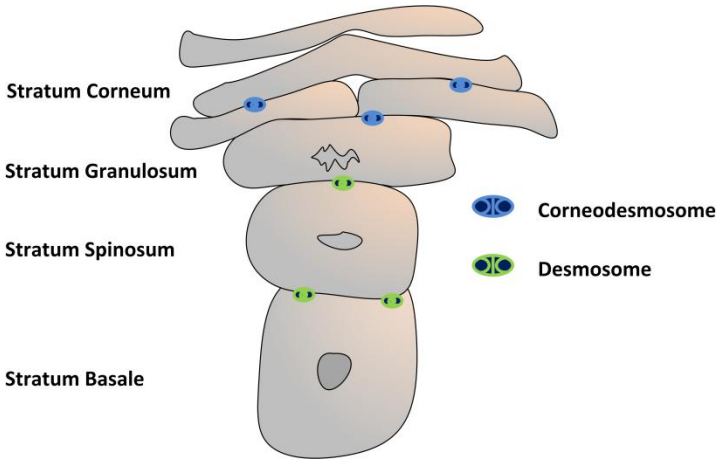

Figure 1-1: Different layers of skin involved in barrier maintenance
Adapted from (Cork et al., 2006)

The epithelium consists of layers of stratum corneum (SC) on the outer side followed by stratum granulosum (SG), stratum spinosum and stratum basale on the inside when observed transversely. The barrier for irritants and allergens is located in the SC. In humans the SC has an average of 20 corneocyte layers, with each corneocyte being approximately 30 µm in diameter**.** Corneocytes are flattened anucleate, organelle-less cells consisting of densely packed keratin fibers. These cells represent the final differentiated stage of the keratinocytes originating from the granular layer (Cork *et al.*,

2006). The thickness of the SC varies within body regions, being thicker to protect areas subjected to greater friction (Lee and Hwang, 2002).

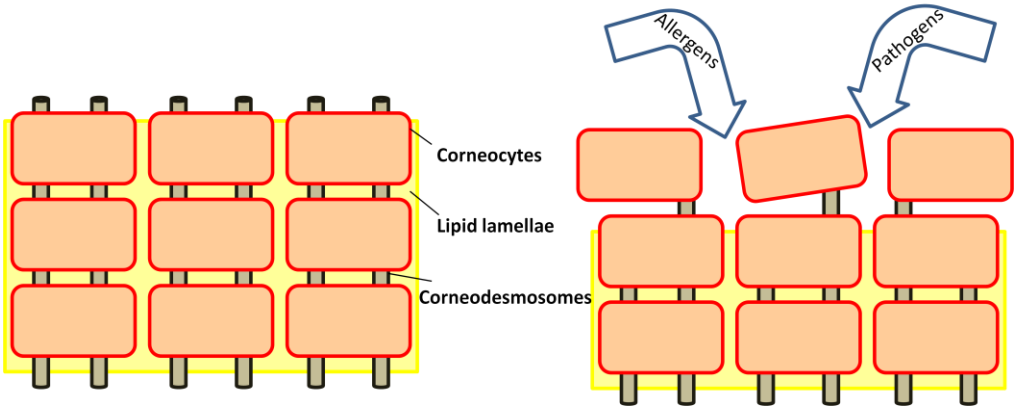

Figure 1-2: Brick wall model of the stratum corneum.
(Adapted from Elias, 1983 & Cork et al, 2006)

As depicted in figure 1.2, the corneocytes are held together by modified desmosomes called corneodesmosomes. Corneodesmosomes help in providing structural integrity and tensile strength to the SC for resisting shearing forces. Elias *et al* visualized the SC as a brick wall, with the corneocytes being the bricks which are cemented by the extra-cellular lipid lamellae (Elias, 1983). Cork *et al* extended this model by envisioning the corneodesmosomes analogous to iron rods that pass down though holes in the bricks to give the wall its tensile strength (Cork *et al.*, 2006). During the formation of corneocytes, the granular cells spill out their lamellar granules by secreting a unique organelle, the epidermal lamellar body (LB), which contains lipid constituents (e.g. cholesterol), lipid precursors (e.g. glucosylceramides and phospholipids), enzymes (e.g. β-glucocerebrosidase, acidic sphingomyelinase,

and secretory phospholipase A2) required for generating ceramides and free fatty acids (FFAs). The lipid lamellae (cement) prevent water loss and penetration of water-soluble materials. Besides, making the barrier as tight as possible they also provide flexibility. The process of keratinocyte differentiation which originates at the stratum basale layer and continues to the corneocyte in the SC is regulated by many proteases and their inhibitors which are also contained in the LB. This cocktail of serine, aspartic and cysteine proteases and their inhibitors orchestrate the orderly digestion of corneodesmosomes and transient intercellular junctions which allows corneocytes to be shed invisibly from the skin surface, a process called desquamation. Desquamation treads a fine balance between adequate breakdown of the barrier to allow a continual renewal of epidermal cells and leaving the barrier sufficiently intact to prevent allergens and irritants from gaining access to the deeper layers of the skin (Elias *et al.*, 2008).

An aberration in squamation and desquamation process can lead to a compromised barrier predisposing an individual to AD (Figure 1.2). Besides a dysregulated physiological processes of skin/ barrier maintenance, sustained psychologic stress also aggravates the permeability barrier dysfunction. **Psychological stress** is both a well-known cause of AD and a cause of resistance to therapy. In experimental animals psychologic stress induces an increase in endogenous glucocorticoids which inhibit the synthesis of the three key epidermal lipids that mediate barrier function, i.e., ceramides, cholesterol, and FFAs (Elias *et al.*, 2008).

The skin's barrier is further safeguarded by a variety of antimicrobial agents and peptides secreted either by the lamellar bodies or in the sweat glands. A

host of antimicrobial peptides (AMPs) have been identified on the skin such as β-defensins, cathelicidins, dermcidins, lactotransferrin, lysozymes, RNase7, amongst others that are permeability increasing proteins thereby exerting bactericidal effect (Schauber and Gallo, 2009). FFA and sphingosines also exhibit antimicrobial activities. Some of these are either constitutively expressed whereas some are inducible by bacterial infections and cytokines. They have a broad spectrum effect covering gram positive, gram negative bacteria, fungi and viruses. Most of the AMPs also have a pro-inflammatory effect making them an active component of innate and adaptive immunity (Hata and Gallo, 2008).

AD is frequently associated with recurrent bacterial and viral infections with over 90% of lesions and 76% on unaffected skin being colonised by *S. aureus* as compared to around 5% in healthy controls (Schittek *et al.*, 2008). Studies have shown a significant reduced expression levels of AMPs in AD skin as compared to psoriatic skin which is another barrier dysfunctional disorder. The **reduced levels of AMPs** in combination with scratching further compromises the barrier property of the skin thereby allowing *S. aureus* and viruses such as HSV to bind and proliferate (Hata and Gallo, 2008).

1.7 Filaggrin in AD

An important gene, widely and most strongly associated with AD is Filaggrin (*FLG*). FLG is one of twenty-plus proteins which gets incorporated in the tough, insoluble layer of the corneocytes (Candi *et al.*, 2005; Oregan *et al.*, 2008). It plays a role in aggregating keratin fibers of the cellular cytoskeleton into bundles, thereby collapsing the corneocytes into flattened discs with a

large surface area. FLG is also extensively deaminated through the actions of the enzyme peptidyl deiminase and subsequently degraded into free amino acids which are then catabolized into the constituents of natural moisturizing factor (NMF), such as lactic acid, sodium pyrrolidone carboxylic acid, urocanic acid, and urea. NMFs are essential for water retention within corneocytes resulting in their optimal hydration and swelling, which prevents the development of gaps thereby enhancing the integrity of the SC and making it resistant to the penetration of irritants and allergens (Elias *et al.*, 2008).

Although initially widely studied in association with ichthyosis vulgaris (IV), it's association to AD was shown in a subset of the IV families and later validated in different European populations (Palmer *et al.*, 2006). Of the many mutations reported, two most common loss-of-function mutations (R501X and 2282del4) account for the majority of AD cases. Up to 50% of AD individuals with European ancestry reveal single- or double-allele or compound mutations in *FLG* (Elias *et al.*, 2008). Subsequently, *FLG*'s association was validated in many different studies such as in a German study on 3000 children (P=2.5×10^{-14}, odds ratio 3.12) (Weidinger *et al.*, 2008). A 1-bp deletion 5303delA in mouse *Flg* analogous to common human *FLG* mutations results in increased allergen priming, cutaneous inflammation and elevated IgE levels against the test allergen, supporting the barrier dysfunction hypothesis (Fallon *et al.*, 2009). In the Japanese population as well, *FLG* has been shown associated with AD, however, a different profile of *FLG* markers seems to exist in different populations (Osawa *et al.*, 2010). FLG's wider variant landscape was further highlighted with an exhaustive discovery search in the Singapore Chinese population where 22 null mutations, 14 of which were

novel, were discovered. Different combination of these 22 null mutations were reported to be associated with AD (Chen *et al.*, 2011).

1.8 AD: a complex disease

AD is a polygenic disease of low penetrance but has a higher prevalence than other monogenic diseases. Similar to other non-Mendelian diseases, AD is multi-factorial with several different (atopic and barrier dysfunction) genetic mechanisms interacting with different environmental factors in different people (Kluken *et al.*, 2003).

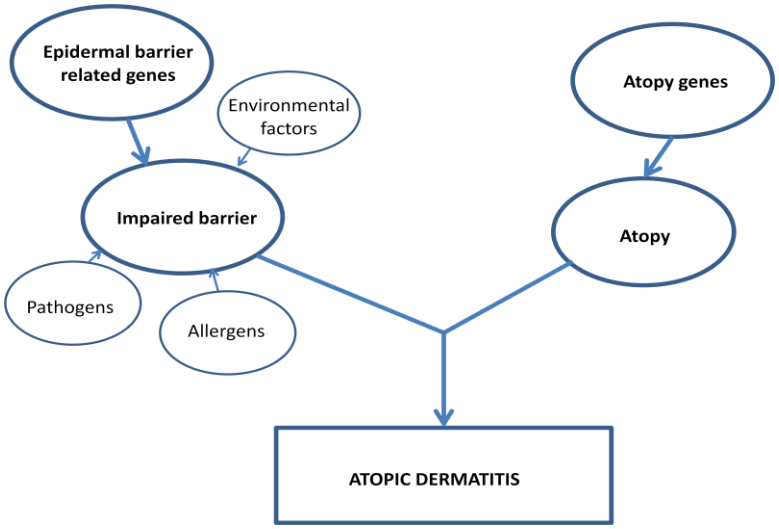

Figure 1-3: Working hypothesis of AD pathogenesis.

With a significant socio-economic impact and increasing prevalence, AD is a viable research candidate. The variability in genes reported for different populations under the influence of different environments makes researching AD in a niche-like setting such as that of Singapore all the more crucial for its better understanding.

1.9 Objectives

With this background information and the working hypothesis of AD manifestation as shown in Figure 1.2, following are the aims of this thesis:

1. To conduct an epidemiological study for identifying and evaluating risk factors associated with eczema in Singapore.
2. To screen and validate barrier maintaining genes associated with AD in ethnic Chinese of Singapore.
3. Using an associated gene, evaluate the effect of the risk variant on gene function.
4. Replicate associations of other relevant studies in Singapore Chinese population.

Chapter 2

Epidemiology of eczema in Singapore

2.1 Introduction

With the need for better understanding the global epidemiology of allergic diseases, International Study of Asthma and Allergies in Childhood (ISAAC) investigated the worldwide prevalence with their standardised questionnaire. This massive effort was undertaken in over 56 countries across multiple centres in each country. With their phase one, they found a wide variation in prevalence of asthma, allergic rhinitis and eczema across the globe (Committee, 1998). The wide variation further highlighted that regional differences, accounted for by different genetic background and environment, contribute to the disease phenotype. Thus, for a deeper understanding of the disease by generating new hypotheses and evaluating the existing hypotheses, a systematic epidemiological analyses in a given geographical context becomes crucial.

In this chapter, I will be describing the study undertaken to evaluate the risk factors for eczema in Singapore in an epidemiological setting.

2.2 Materials and methods:

2.2.1 Population ascertainment

Population was ascertained using a cross-sectional sampling methodology from National University of Singapore's (NUS) campus. The recruitment drives targeted undergraduate and graduate students and staff from all faculties on campus above 18years of age. The sample recruitment drives were organised every year for over seven years and is still being continued. The recruitment drives were approved by the Institutional Review Board (NUS-IRB Reference Code: 07-023 and NUS10-343) and were in accordance with the Helsinki Declaration. All participants recruited in NUS were compensated for their participation in the study with S$10 vouchers.

Participants were also tested for their atopic status by performing a skin prick test (SPT). The participants were tested for four aeroallergens namely *Dermatophagoides pteronyssinus, Blomia tropicalis, Elaeis guineensis* and *Curvularia lunata*. To validate the test, a positive and negative control i.e. histamine and saline (buffer used for the preparation of the allergen) respectively were also skin pricked. SPT was considered positive when the wheal diameter was ≥ 3mm whereas no signs of wheal and erythema was considered a negative SPT. Participants with less than 3mm reaction for histamine or, an erythema or wheal to saline, or smaller than 3mm reaction to the test allergens were not considered for analyses. Studies done by the lab previously showed that skin-prick reactivity to dust mites especially the two species *Blomia tropicalis* and *Dermatophagoides pteronyssinus* were highly sensitive and specific markers for allergic sensitization in Singapore (Chew *et al.*, 1999). Thus, together with the positive and negative controls, a positive reaction to either of the dust mites indicated positive SPT or atopy.

2.2.2 Classification and disease definitions

For disease classification participants filled out the standardised and validated ISAAC questionnaire (Asher *et al.*, 1995). Core questions of ISAAC provided information on current (last 1 year) and past symptoms for asthma, allergic rhinitis (AR) and eczema. The questionnaire also had sections for demographic, medical, dietary, lifestyle and family information. Additional question from other UK working party criteria (Williams *et al.*, 1994a) on dry skin was added at a later stage of the collection. The complete questionnaire is attached in the appendix.

A participant was categorised as a case for eczema if he or she answered affirmative to "ever had itchy rash" and "rash at skin flexures". Based on UK working party's criteria, flexural involvement of rashes had 90% sensitivity and 75% specificity for diagnosing AD (Williams *et al.*, 1994a).

2.2.3 Associations with eczema

All the queried independent variables in the questionnaire were test for association to eczema. The list was exhaustive based on the questionnaire. Categorical variables were analysed for the dichotomous disease outcome. Continuous variables such as age and number of years in Singapore were split into groups and coded categorically. Associations were checked individually for each independent variable giving odds ratio (OR) and their respective 95% confidence interval (95% CI) and were considered significant if the p value was less than 0.05 and odds ratio range did not cross 1. Independent variables were also re-tested adjusting for associated demographic factors such as gender, age group, race. This was done to control for any false positives arising due to the confounding effect of demographics on disease outcome.

2.2.4 Statistical Analyses:

All data entry and statistical analyses were done using Microsoft Excel (http://office.microsoft.com/en-us/excel/) and SPSS (http://www.ibm.com/software/analytics/spss/). The binary logistic regression function from SPSS was used for calculating the crude and adjusted odds ratio.

2.3 Results

2.3.1 Population description

As the collection was at the university campus, we obtained a large diversity in ages participating. However, for this analysis the age groups considered is from 18 to 32, subjects above 32 years of age were excluded. After excluding samples not matching our age band, we obtained 8334 samples as shown in Table 2.1. The male to female ratio were almost similar. The age groupings were done to avoid any skews towards a group with around 36% of the population between the 18-20 and 20.01-22 age groups. The pre-dominance of ethnic Chinese in Singapore's population resulted in 91% of samples recruited to be ethnic Chinese, followed by 4% Indians, 2% Malays and the remaining were either of mixed heritage or from other races. As shown in Table 2.1, close to 63% lived in flats provided at subsidy by the Singapore government, around 17% in condominiums and around 13% in landed property. Monthly family income in Singapore dollars did not reveal a wide difference between the 4 classes. Singapore's diverse background is evident by the proportion of participants born in countries other than Singapore. The international student

community accounted for the 20% of population living in Singapore for less than 5 years.

Variables		Number	(%)
Gender			
	Male	3615	(43.38)
	Female	4707	(56.48)
	Not stated	12	(0.14)
Age group			
	18-20	3074	(36.89)
	20.01-22	3048	(36.57)
	22.01-32	2212	(26.54)
Racial groups			
	Chinese	7565	(90.77)
	Malay	158	(1.90)
	Indians	331	(3.97)
	Others	280	(3.36)
Housing			
	Public flats	5261	(63.13)
	Private apartments	1433	(17.19)
	Landed property	1137	(13.64)
	Not stated	503	(6.04)
Family income (S$/month)			
	Less than 2000	2075	(24.90)
	2000-3999	2858	(34.29)
	4000-5999	1493	(17.91)
	6000 and above	1579	(18.95)
	Not stated	329	(3.95)
Country of birth			
	Singapore	5517	(66.20)
	Others	2781	(33.37)
	Not stated	36	(0.43)
No. of years in SG			
	0-5	1691	(20.29)
	5.01-10	420	(5.04)
	10.01-32	5891	(70.69)
	Not stated	332	(3.98)
Total studied		**8334**	**(100)**

Table 2-1: Demographic profile and socioeconomic categories of study population.

Abbreviations: S$- Singapore dollar, SG- Singapore, No.-number

2.3.2 Self-reported prevalence of asthma, allergic rhinitis and eczema

The response rates for all three-disease sections of the questionnaire were high with asthma section being responded by 100%, allergic rhinitis (AR) by 99.74% and AD by 99.72% of the participants (Table 2.2).

Five percent of the asthma respondents reported wheezing in the last one year with 18% reporting in positive to the question "ever had wheezing?" indicating that most of them had childhood asthma. The response to the question "ever had wheezing?" (‘ever wheezing’) had a higher positive response from males as compared to females and it correlated well with self diagnosed asthma. Self diagnosed asthma was queried by the question 'have you ever had asthma' coded as 'ever asthma'. Stratification based on races indicated that Indians followed by Malays showed higher prevalence of wheezing in the last one year and wheezing during exercising. However, only 16% of Indians self-reported asthma as compared to 23% Malays indicating lack of awareness. With respect to family income, a trend of increasing prevalence of asthma symptoms was observed with increasing family income. It was also found that wheezing and self reported asthma were reported higher amongst people born in Singapore.

The trend of growing out of allergic diseases wasn't observed in AR with 51% reporting AR in the past one-year and a total of 56% reporting in positive to ever having it. When stratified with respect to gender it was seen that males had a higher percentage of self-reporting AR as compared to females. Allergic rhinitis had an equal spread amongst the races except for the higher prevalence of self reported AR in the Chinese as compared to Malays. It was also observed that the people born in Singapore and those who had spent 10.01-32

years in Singapore had a higher prevalence of AR within the last one year as compared to people born in other countries.

The percentage of participants reporting in positive to ever having a rash was similarly distributed to the ones who had a chronic rash in the past one year. Eighteen percent of the respondents claimed to have flexural rash which is a hallmark of AD, whereas only 13% reported the presence of eczema indicating an under diagnosis of the disease. With respect to gender, females reported a higher incidence of chronic rash at flexures. Race stratification showed that as compared to the Indians and Chinese the Malays had the highest prevalence for eczema related symptoms. It was also observed that people born in Singapore and those who had spent 10.01-32 years in Singapore had a higher percentage of self-reporting eczema as compared to people born in other countries.

Symptoms	Overall	Gender		Age grp			Race			Housing			Family income (S$/month)				Country of birth		No of years in Singapore		
		Male	Female	18-20	20.01-22	22.01-32	Chinese	Malay	Indian	Public	Private	Landed	< 2K	2-3.999K	4-5.999K	>=6K	Singapore	Other	0-5	5.01-10	10.01-32
Asthma responses (N)	**8334**	**3615**	**4707**	**3074**	**3048**	**2212**	**7565**	**158**	**331**	**5261**	**1433**	**1137**	**2075**	**2858**	**1493**	**1579**	**5517**	**2781**	**1691**	**420**	**5891**
(%)	100	43.38	56.48	36.89	36.57	26.54	90.77	1.90	3.97	63.13	17.19	13.64	24.90	34.29	17.91	18.95	66.20	33.37	20.29	5.04	70.69
Ever wheezed (%)	18.49	20.41	16.97	18.12	19.88	17.09	17.90	25.32	25.68	18.29	19.75	20.14	14.12	18.82	20.03	23.05	20.63	14.20	14.13	11.19	20.54
Whz in last year (%)	4.93	4.90	4.97	5.01	5.18	4.48	4.47	6.96	10.57	5.06	4.54	5.28	4.24	5.14	5.02	5.76	5.40	3.96	4.49	2.38	5.36
Whz when exercise (%)	3.64	3.07	4.04	4.20	3.58	2.94	3.19	5.06	8.46	3.23	3.84	4.66	3.42	3.18	4.02	4.43	3.30	4.28	5.09	2.62	3.33
Cough at night (%)	8.63	7.75	9.28	9.50	9.22	6.60	8.17	7.59	13.90	8.10	9.42	10.03	8.19	8.15	9.18	9.44	8.48	8.88	9.76	7.62	8.44
Self reported asthma (%)	18.45	21.38	16.19	18.74	20.31	15.51	18.57	22.78	15.71	18.53	18.77	20.40	14.02	18.61	20.76	22.61	21.75	11.90	11.06	12.14	21.49
Rhinitis responses (N)	**8312**	**3604**	**4696**	**3063**	**3041**	**2208**	**7547**	**158**	**328**	**5247**	**1427**	**1136**	**2068**	**2849**	**1490**	**1576**	**5503**	**2774**	**1686**	**419**	**5877**
(%)	100	43.36	56.50	36.85	36.59	26.56	90.80	1.90	3.95	63.13	17.17	13.67	24.88	34.28	17.93	18.96	66.21	33.37	20.28	5.04	70.71
Ever rhinitis (%)	56.27	57.88	55.03	55.63	56.69	56.57	56.30	60.13	49.70	56.85	53.75	60.04	52.51	57.46	57.05	59.58	58.86	51.30	51.01	46.78	58.84
Rhinitis in last year (%)	51.44	53.19	50.11	50.28	52.25	51.95	51.45	53.80	46.65	52.32	48.77	54.14	47.49	53.21	51.95	54.25	54.15	46.18	45.79	42.96	54.13
Rhinoconjunctivitis (%)	20.99	20.34	21.53	20.08	21.60	21.42	20.43	22.15	27.44	21.29	19.97	22.27	18.18	21.59	23.02	22.40	22.77	17.48	17.02	16.95	22.65
Self reported AR (%)	6.58	8.10	5.41	4.24	7.56	8.47	6.47	3.80	5.49	6.17	7.22	7.92	6.14	6.35	6.64	7.80	6.07	7.57	7.65	6.44	6.26
Eczema responses (N)	**8311**	**3605**	**4694**	**3062**	**3040**	**2209**	**7546**	**158**	**328**	**5253**	**1427**	**1132**	**2070**	**2854**	**1485**	**1575**	**5512**	**2774**	**1684**	**419**	**5887**
(%)	100	43.38	56.48	36.84	36.58	26.58	90.80	1.90	3.95	63.21	17.17	13.62	24.91	34.34	17.87	18.95	66.32	33.38	20.26	5.04	70.83
Ever chronic rash (%)	20.98	19.17	22.37	20.80	19.57	23.18	20.99	25.95	17.68	20.58	22.28	21.20	20.24	21.13	21.28	21.46	21.37	20.04	20.84	21.96	21.06
Chrn rash last year (%)	20.61	18.22	22.41	20.25	19.24	23.00	20.46	28.48	16.46	20.18	21.86	20.05	20.72	20.67	20.74	20.32	20.34	20.98	22.03	21.96	20.18
Flexural rash (%)	18.28	16.39	19.71	18.03	17.76	19.33	18.21	25.32	14.02	18.16	18.71	17.67	17.78	18.96	17.98	17.65	18.54	17.56	18.47	21.96	18.31
Self reported eczema (%)	13.78	12.76	14.53	14.01	14.21	12.86	14.09	13.92	7.01	13.27	15.49	15.02	10.14	12.68	16.90	17.14	15.31	10.71	10.27	21.96	15.20

Table 2-2: Self-reported prevalence of asthma, rhinitis and eczema and their related symptoms.

Abbreviations: Whz- Wheeze, Chrn- Chronic, grp- group, S$- Singapore dollar, K- 1000S, N, No.- number

2.3.3 Self-reported prevalence of *severity* of asthma, allergic rhinitis and eczema.

Severity of the three allergic diseases was evaluated for those who reported symptoms in the last one year (Table 2.3).

Of the asthma respondents having wheezing in the last one year, 70% of them had 1-3 wheezing episodes, 21% had 4-12 episodes and about 5% had more than 12 episodes of wheezing averaging more than once a month in the past year. While the majority of the current wheezers seemed to have mild asthma not causing sleep disturbance (61%), 10.5% suffered wheezing strong enough to be woken up form sleep more than 1 night per week and close to 12% reporting speech limitation due to wheezing. The stratification based on gender showed no particular bias with the exception of speech limitation wherein females reported twice as more prevalent than in males (8%). Based on age group our data showed a higher prevalence (27%) of 4-12 wheezing attacks in the 20.01-22 age band. The oldest age group reported higher prevalence of no interference with daily activities as compared to the other two age groups. Unfortunately, the stratification based on race for asthma severity cannot be commented as only 11 Malays reported wheezing in the last one year and percentages calculated off this number could be unreliable. Based on housing our data showed a trend of higher prevalence of severity of asthma in private housing as compared to public or landed. Symptoms such as higher frequency of wheezing attacks and speech limitation did show a trend of increased prevalence in the more affluent families (15% versus 7%).

Amongst the sufferers with AR symptoms in the past one-year, majority (37%) reported mild interference with daily activities and close to 3% reported

severe interference. Twenty-five percent found their AR symptoms troublesome with 22% reporting sleep disturbance. Largest proportion (29%) has been living with AR symptoms since more than 10 years. Of these a higher proportion of males are reported to be living with AR symptoms for more than 10 years (33%). Based on race, Indians reported to have higher prevalence of interference of daily activities and disturbances due to AR symptoms. Severity of rhinitis did not seem to fluctuate much, when binned according to type of housing. Also the severity symptoms showed no changes when compared across the four income groups. People living with AR symptoms for the past 1-4 years was highest amongst the group of people staying in Singapore for 5.01-10 years (31%) followed by 0-5 years (24%) and was least in people staying for more than 10 years (17%).

Close to 50% of the respondents claiming chronic rash in the past one year reported the rash occurring first above the age of 5. Thirty-four percent of them reported persistence rash with the rash not completely clearing up in the past one year. Severity measured by being kept awake due to rash showed 60% of mild form to 10% of the severe form. Question on dry skin was incorporated in the questionnaire much later in the collection years. This resulted in the number of respondents with chronic rash, which encountered the question on dry skin as 793. Of the 793 individuals, 52% reported to have dry skin. Of these females were reported higher at 57% as compared to 43% males. A trend of decreasing severity of eczema was reported with the increase in age. Stratification based on race showed that Malays have a higher prevalence of persistent rash in the past year followed by Chinese, with Indians reporting the least prevalence of such a persistent rash. Malays again

reported a higher prevalence of dry skin as compared to the other two races. The data showed no variations depending on the type of housing, but showed a higher prevalence of early onset eczema and more severe symptoms disturbing sleep in the highest income slab group. People living in Singapore for more than 10 years had a higher prevalence of dry skin (57%) and more severe form of eczema (11%) as compared to those staying for less than 5 years (43% and 5% respectively).

The gender ratio for asthma and AR were similar, however, there were close to twice the number of females who reported eczema in the last 1 year. The higher age group of 22.01- 32 years showed to have a lower prevalence for all three diseases. The racial distribution of all the diseases was similar to the racial distribution in the population with the exception of Indians having higher self-reported wheezing in the last one year. The distribution with regards to the type of housing, family income, country of birth and years in Singapore for all three diseases was similar to the general populations distribution.

Symptoms	Overall	Gender		Age grp			Race			Housing			Family income (S$/month)				COB		No of years in S'pore		
		Male	Female	18-20	20.01-22	22.01-32	Chinese	Malay	Indian	Public	Private	Landed	< 2K	2-3.9K	4-5.9K	>=6K	S'pore	Other	0-5	5.01-10	10.01-32
Wheezing in the last year (N)	**411**	**177**	**234**	**154**	**158**	**99**	**338**	**11**	**35**	**266**	**65**	**60**	**88**	**147**	**75**	**91**	**298**	**110**	**76**	**10**	**316**
(%)	100	43.1	56.9	37.5	38.4	24.1	82.2	2.7	8.5	64.7	15.8	14.6	21.4	35.8	18.2	22.1	72.5	26.8	18.5	2.4	76.9
No. Of wheezing episodes (%)																					
1-3	69.8	68.9	70.5	72.7	64.6	73.7	68.9	90.9	68.6	70.3	64.6	68.3	68.2	70.7	70.7	69.2	71.8	64.5	65.8	70.0	70.6
4-12	21.4	22.0	20.9	18.2	27.2	17.2	21.6	9.1	22.9	21.8	21.5	21.7	21.6	23.1	18.7	19.8	20.5	23.6	25.0	20.0	20.9
>12	5.1	5.1	5.1	4.5	5.1	6.1	5.6	0.0	2.9	3.8	10.8	6.7	3.4	2.7	9.3	7.7	5.4	4.5	3.9	0.0	5.7
Woken by wheeeze (%)																					
Nil	61.3	62.1	60.7	61.7	62.0	59.6	61.2	63.6	54.3	63.5	50.8	58.3	64.8	59.2	65.3	59.3	63.4	55.5	55.3	50.0	63.6
<1 per week	26.3	24.9	27.4	26.0	25.3	28.3	25.7	27.3	31.4	23.7	35.4	28.3	22.7	27.2	24.0	29.7	24.8	30.0	32.9	30.0	24.4
>1 per week	9.5	9.6	9.4	9.1	10.1	9.1	9.8	9.1	11.4	9.4	12.3	10.0	8.0	10.9	9.3	7.7	9.4	10.0	9.2	10.0	9.5
Speech limitation (%)	12.2	7.9	15.4	13.6	12.7	9.1	11.8	18.2	8.6	11.7	18.5	11.7	8.0	11.6	10.7	15.4	12.8	10.9	9.2	10.0	13.0
Rhinitis in the last year (N)	**4276**	**1917**	**2353**	**1540**	**1589**	**1147**	**3883**	**85**	**153**	**2745**	**696**	**615**	**982**	**1516**	**774**	**855**	**2980**	**1281**	**772**	**180**	**3181**
(%)	100	44.8	55.0	36.0	37.2	26.8	90.8	2.0	3.6	64.2	16.3	14.4	23.0	35.5	18.1	20.0	69.7	30.0	18.1	4.2	74.4
Intrfrnc with daily activity (%)																					
None	22.8	24.4	21.5	17.7	17.9	36.6	23.5	23.5	13.1	23.8	19.4	22.1	22.4	24.1	22.0	20.2	22.9	22.8	19.0	21.7	22.9
Mild	37.7	34.7	40.1	37.6	37.8	37.8	37.8	37.6	35.9	37.6	38.1	39.5	39.0	36.8	37.0	40.7	37.3	38.7	39.6	40.0	37.4
Moderate	28.7	30.7	27.1	28.4	33.7	22.2	28.1	24.7	37.9	28.7	33.8	26.0	25.8	28.6	32.2	29.7	29.3	27.4	28.9	29.4	29.2
Severe	3.0	2.9	3.1	3.1	3.1	2.5	2.8	3.5	5.9	3.0	3.2	3.3	2.2	3.0	3.2	3.6	2.8	3.4	3.5	2.8	2.9
Disturbances (%)																					
Sleep disturbance	21.9	19.2	24.0	23.6	21.3	20.4	21.4	15.3	28.8	20.9	25.3	24.1	23.5	20.1	20.9	24.3	20.8	24.6	25.1	28.3	21.0
Impairment of leisure/sport	15.9	17.2	14.9	14.8	16.0	17.3	15.2	11.8	26.8	15.2	18.4	17.4	15.3	17.7	15.1	15.7	14.5	19.3	20.7	18.3	14.7
Impairment of school/work	14.0	14.9	13.2	13.5	13.7	15.0	13.4	16.5	20.3	13.4	14.7	15.8	13.8	13.9	15.5	13.5	13.6	14.9	16.6	12.2	13.5
Troublesome symptoms	25.7	28.7	23.2	22.9	27.4	27.2	26.1	15.3	21.6	26.1	29.2	23.6	22.4	25.4	27.9	29.4	27.5	21.3	19.9	20.6	27.8
Duration of nose symptoms (%)																					
<1 yr	10.5	8.3	12.2	11.2	10.4	9.8	10.1	17.6	13.7	10.7	10.2	10.4	10.8	11.0	9.2	11.1	9.8	12.3	13.2	12.2	9.9
1-4 yrs	18.6	17.9	19.2	19.9	18.6	17.0	18.2	17.6	22.9	18.1	19.7	20.2	20.4	19.1	17.2	17.4	16.5	23.7	24.5	30.6	16.7
5-10yrs	16.0	15.4	16.4	16.4	15.7	15.8	15.9	9.4	17.0	16.1	18.0	14.1	16.0	14.9	17.4	16.7	16.1	15.7	14.9	14.4	16.4
>10 years	28.8	33.0	25.4	23.4	33.4	29.8	29.4	16.5	30.1	29.1	32.6	28.3	24.9	28.2	32.0	33.5	31.3	23.0	23.2	22.2	31.0
Eczema in the last year (N)	**1713**	**657**	**1052**	**620**	**585**	**508**	**1544**	**45**	**54**	**1060**	**312**	**227**	**429**	**590**	**308**	**320**	**1121**	**582**	**371**	**95**	**1188**
(%)	100	38.4	61.4	36.2	34.2	29.7	90.1	2.6	3.2	61.9	18.2	13.3	25.0	34.4	18.0	18.7	65.4	34.0	21.7	5.5	69.4
Age of onset (%)																					
<2yrs	6.8	7.5	6.5	6.5	8.0	5.9	6.7	13.3	3.7	5.4	11.9	8.8	4.9	6.3	5.5	12.2	7.0	6.4	6.7	5.3	7.2
2-4yrs	5.3	5.3	5.3	6.5	4.1	5.3	5.3	4.4	1.9	4.9	6.4	6.2	4.0	5.1	6.2	6.9	6.4	3.3	3.8	2.1	6.1
5 and above	49.4	45.8	51.4	49.7	48.5	50.0	50.0	31.1	55.6	51.1	47.4	42.7	48.7	53.4	47.1	47.8	49.7	49.3	49.3	48.4	49.3
Persistent rash (%)	34.3	33.8	34.7	37.6	35.0	29.5	34.7	53.3	16.7	35.3	36.2	32.2	35.0	33.1	36.0	34.7	35.8	31.3	32.6	30.5	35.4
Kept awake by rash (%)																					
Never	59.5	62.4	57.6	56.3	57.9	65.4	59.2	51.1	66.7	59.8	54.2	62.6	62.0	58.5	60.4	56.6	56.5	65.3	69.5	55.8	56.9
<1 night per week	28.5	25.6	30.4	31.3	29.1	24.4	28.6	35.6	24.1	29.2	31.1	24.2	26.3	30.7	26.3	29.1	31.2	23.4	22.9	27.4	30.4
>1 night per week	9.7	9.4	10.0	10.5	10.1	8.5	9.9	13.3	7.4	9.2	11.5	9.7	8.6	8.8	12.0	12.5	10.6	8.2	5.4	9.5	10.9
Dry Skin (%)*	52.2	44.9	56.6	51.7	55.2	47.9	50.8	75.0	57.1	53.2	56.2	51.9	46.7	49.2	59.4	57.0	57.6	43.8	43.5	44.6	57.2

Table 2-3: Self-reported prevalence of severity symptoms for asthma, rhinitis and eczema.

*Abbreviations: Intrfrnc-Interference grp- group, S$- Singapore dollar, K- 1000$, N, No.- number, S'pore- Singapore, COB-Country of birth * Dry skin calculated out of 793 instead of 1713*

2.3.4 Association of covariates

2.3.4.1 Evaluation association of demographics to eczema

Demographic factors such as gender, age, race, housing, monthly family income, country of birth and number of years in Singapore were tested for association to eczema. Eczema was defined as chronic rash having a flexural involvement which made up to 1310 individuals from the 8334 ascertained, bringing its prevalence to 15.72 %. Continuous data such as age and number of years in Singapore were categorised into three groups as shown in the above tables. Logistic regression for outcome of disease was used to obtain Table 2.4.

Table 2.4 shows that females had a higher odds of having eczema as compared to males (OR 1.23). Age group also showed mild association with the more older group having a higher prevalence of eczema (OR1.16) as compared to the other 2 age groups. Race also showed a disparity with Malays having higher odds of developing eczema as compared to Indians. No associations were found at the housing, family income, country of birth and years spent in Singapore to the disease outcome.

As demographics such as gender, age group and race were associated with eczema, analyses for other variables may be confounded by their effect. Hence, the following variables - gender, age and race were used for adjusting the prevalence odds ratio of remaining covariates.

Categorical variables	Odds ratio	(95% CI)	p value
Gender			
Male	Ref	NA	NA
Female	**1.23**	**(1.09-1.39)**	**0.001**
Age group			
20.01-22	Ref	NA	NA
18-20	1.16	(0.94-1.24)	0.271
22.01-32	**1.16**	**(1.00-1.35)**	**0.042**
Race			
Indian	Ref	NA	NA
Chinese	1.39	(0.99-1.96)	0.054
Malay	**1.97**	**(1.18-3.28)**	**0.009**
Housing			
Public	Ref	NA	NA
Private	1.05	(0.90-1.24)	0.485
Landed	1.04	(0.87-1.24)	0.629
Family income (S$/month)			
< 2K	Ref	NA	NA
2-3.999K	0.95	(0.81-1.11)	0.531
4-5.999K	0.95	(0.80-1.13)	0.573
>=6K	0.95	(0.80-1.12)	0.555
Country of Birth			
Others	Ref	NA	NA
Singapore	1.09	(0.96-1.24)	0.159
Yrs in Singapore			
10.01-32	Ref	NA	NA
0-5	1.01	(0.87-1.17)	0.885
5.01-10	0.93	(0.71-1.23)	0.656

Table 2-4: Association of demographic variables to eczema.
Abbreviations: CI- confidence interval, Ref- reference, NA- Not applicable, S$- Singapore dollar, K- $1000, Yrs- years

Variable	Odds ratio	(95% CI)	p value
Gender			
Male	Ref	NA	NA
Female	**1.29**	**(1.129 - 1.478)**	**<0.001**
Age group			
20.01-22	Ref	NA	NA
18-20	1.01	(0.867 - 1.176)	NS
22.01-32	**1.17**	**(1.003 - 1.364)**	**0.046**
Race			
Indian	Ref	NA	NA
Chinese	**1.45**	**(1.031 - 2.040)**	**0.033**
Malay	**2.06**	**(1.237 - 3.432)**	**0.005**
Constant	0.1	NA	<0.001

Table 2-5: Demographic variables used for adjusting odds ratio.

Categorical variables	Odds ratio	(95% CI)	p value	Adjusted OR	(95% CI)	p value
Housing						
Public	Ref	NA	NA	Ref	NA	NA
Private	1.05	(0.90-1.24)	0.485	1.11	(0.94 - 1.31)	0.192
Landed	1.04	(0.87-1.24)	0.629	1.07	(0.89 - 1.29)	0.417
Family income (S$/month)						
< 2K	Ref	NA	NA	Ref	NA	NA
2-3.999K	0.95	(0.81-1.11)	0.531	1.07	(0.91 - 1.26)	0.363
4-5.999K	0.95	(0.80-1.13)	0.573	1.03	(0.85 - 1.24)	0.727
>=6K	0.95	(0.80-1.12)	0.555	1.04	(0.87 - 1.26)	0.612
Country of Birth						
Others	Ref	NA	NA	Ref	NA	NA
Singapore	1.09	(0.96-1.24)	0.159	1.12	(0.98 - 1.28)	0.084
Yrs in Singapore						
10.01-32	Ref	NA	NA	Ref	NA	NA
0-5	1.01	(0.87-1.17)	0.885	1	(0.85 - 1.17)	0.957
5.01-10	0.93	(0.71-1.23)	0.656	0.93	(0.70 - 1.24)	0.649

Table 2-6: Demographics' prevalence odds ratio for eczema. Uni-variate and multivariate analyses showing crude and adjusted odds ratio.
Abbreviations: CI- confidence interval, Ref- reference, NA- Not applicable, S$- Singapore dollar, K- $1000, Yrs- years

As shown in the table above, the three variables remain significant even when controlling for the presence of the other covariates. Interestingly, when

controlling for gender and age, prevalence odds for Malays increases slightly and Chinese ethnicity also gets highlighted as a race with significantly higher eczema outcome as compared to Indians.

The demographics earlier found to be not significant were re-evaluated by adjusting for the model (Table 2.5) to check for any confounding effect (Table 2.6).

After adjusting for the variables, the other demographic variables did not show any association.

2.3.4.2 Evaluating association of diet and lifestyle to eczema

Questions on diet were also collected and checked for association. Unadjusted and adjusted odds ratio for the different diet choices are shown in table 2.7 below.

Prevalence of eczema was higher amongst people consuming pulses, nuts, potatoes, fast-food and yogurt drinks on most days as compared to people consuming it once-to-twice a week. Compared to daily consumption, once-to-twice a week intake of cereals and rice had a higher prevalence of eczema. Participants with occasional consumption of margarine had a higher prevalence of eczema as compared to once-to-twice a week.

Categorical variables	Odds ratio	(95%CI)	p value	Adjusted OR	(95% CI)	p value
Meat						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.17	(0.82 - 1.66)	0.374	1.32	(0.91 - 1.90)	0.135
Most/ all days	1.01	(0.84 - 1.23)	0.855	1	(0.82 - 1.22)	0.951
Seafood						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.14	(0.92 - 1.41)	0.202	1.21	(0.97 - 1.52)	0.084
Most/ all days	1.01	(0.90 - 1.15)	0.761	1.02	(0.90 - 1.16)	0.695
Fruits						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.05	(0.77 - 1.44)	0.73	1	(0.72 - 1.39)	0.970
Most/ all days	1.01	(0.89 - 1.14)	0.872	0.97	(0.85 - 1.10)	0.661
Vegetables						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.24	(0.84 - 1.83)	0.264	1.14	(0.76 - 1.71)	0.515
Most/ all days	1.01	(0.84 - 1.21)	0.89	1	(0.83 - 1.20)	0.955
Pulses						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.04	(0.89 - 1.22)	0.565	1.03	(0.88 - 1.21)	0.665
Most/ all days	1.11	(0.96 - 1.29)	0.136	**1.17**	**(1.01 - 1.36)**	**0.038**
Cereals						
Most/ all days	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.03	(0.83 - 1.28)	0.727	1.05	(0.84 - 1.32)	0.634
Once or twice/week	**1.16**	**(1.02 - 1.31)**	**0.018**	**1.2**	**(1.05 - 1.36)**	**0.005**
Pasta						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.04	(0.92 - 1.18)	0.494	1.02	(0.89 - 1.16)	0.757
Most/ all days	1.14	(0.93 - 1.39)	0.189	1.12	(0.91 - 1.38)	0.256
Rice						
Most/ all days	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.1	(0.72 - 1.68)	0.635	0.06	(0.61 - 1.49)	0.86
Once or twice/week	**1.24**	**(1.02 - 1.50)**	**0.024**	**1.23**	**(1.01 - 1.50)**	**0.036**
Butter						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	0.99	(0.87 - 1.12)	0.904	0.95	(0.83 - 1.08)	0.460
Most/ all days	0.95	(0.78 - 1.16)	0.672	0.99	(0.81 - 1.21)	0.951
Margarine						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	**1.18**	**(1.04 - 1.34)**	**0.01**	**1.16**	**(1.02 - 1.32)**	**0.024**
Most/ all days	1.05	(0.84 - 1.31)	0.648	1.07	(0.86 - 1.34)	0.512
Nuts						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.08	(0.96 - 1.23)	0.179	1.05	(0.92 - 1.19)	0.436
Most/ all days	**1.25**	**(1.00 - 1.56)**	**0.041**	**1.29**	**(1.03 - 1.62)**	**0.024**
Potatoes						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.2	(1.03 - 1.40)	0.022	1.15	(0.98 - 1.35)	0.084
Most/ all days	1.16	(0.99 - 1.36)	0.064	**1.21**	**(1.03 - 1.43)**	**0.021**
Milk						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.09	(0.94 - 1.28)	0.235	1.08	(0.92 - 1.27)	0.317
Most/ all days	1.01	(0.88 - 1.15)	0.854	1.02	(0.89 - 1.18)	0.709
Eggs						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.08	(0.80 - 1.45)	0.59	1.04	(0.75 - 1.43)	0.786
Most/ all days	1.05	(0.93 - 1.19)	0.369	1.09	(0.97 - 1.24)	0.138
Burgers/ fast food						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1.08	(0.95 - 1.22)	0.22	1.02	(0.90 - 1.17)	0.679
Most/ all days	**1.36**	**(1.08 - 1.70)**	**0.008**	**1.37**	**(1.09 - 1.73)**	**0.007**
Yogurt drinks						
Once or twice/week	Ref	NA	NA	Ref	NA	NA
Never/occasionally	1	(0.88 - 1.13)	0.984	0.99	(0.87 - 1.14)	0.989
Most/ all days	**1.19**	**(1.001 - 1.42)**	**0.049**	**1.2**	**(1.00 7- 1.44)**	**0.041**

Table 2-7: Diet habit's prevalence odds ratio for eczema. Uni-variate and multivariate analyses showing crude and adjusted odds ratio.

Questions on lifestyle such as exercise frequency, sedentary lifestyle, alcohol consumption, smoke exposure, pet exposure were asked in the questionnaire. In addition, drug allergy information and allergic sensitisation as evaluated by skin prick test were also analysed in the following table 2.8.

Categorical variables	Odds ratio	(95%CI)	p value	Adjusted OR	(95% CI)	p value
Vigorous physical activity/week						
1-2 times	Ref	NA	NA	Ref	NA	NA
Most/ all days	0.97	(0.80 - 1.19)	0.838	1.01	(0.82 - 1.24)	0.892
Never/occasionally	1.06	(0.93 - 1.20)	0.354	0.99	(0.86 - 1.13)	0.88
TV/ Computer (hrs/day)						
<1	Ref	NA	NA	Ref	NA	NA
1-5	1.08	(0.79 - 1.06)	0.249	1.11	(0.96 - 1.29)	0.132
5<	1.23	(0.95 - 1.34)	0.139	**1.29**	**(1.05 - 1.57)**	**0.012**
Alcohol consumption						
Non-drinker	Ref	NA	NA	Ref	NA	NA
Frequent	1.22	(0.82 - 1.82)	0.323	1.28	(0.82 - 1.98)	0.267
Occasional	0.96	(0.84 - 1.08)	0.52	1	(0.88 - 1.14)	0.911
Smoke exposure						
No	Ref	NA	NA	Ref	NA	NA
Yes	0.95	(0.83 - 1.08)	0.486	0.92	(0.80 - 1.05)	0.244
Ever had pets						
No	Ref	NA	NA	Ref	NA	NA
Yes	1.14	(0.99 - 1.29)	0.075	1.09	(0.95 - 1.26)	0.199
Drug allergies						
No	Ref	NA	NA	Ref	NA	NA
Yes	**1.39**	**(1.16-1.66)**	**<0.001**	**1.45**	**(1.21 - 1.75)**	**<0.001**
Allergic sensitisation						
No	Ref	NA	NA	Ref	NA	NA
Yes	**1.54**	**(1.31-1.81)**	**<0.001**	**1.72**	**(1.45 - 2.05)**	**<0.001**

Table 2-8: Lifestyle, exposure and allergic status evaluated for eczema. Univariate and multivariate analyses showing crude and adjusted odds ratio.

Physical activity, alcohol consumption, smoke exposure, pet exposure were not associated with eczema in this dataset. After adjusting for the three demographic variables, sedentary lifestyle of sitting in front of a computer or television for more than 5 hours had a higher prevalence of eczema as compared to less than 1 hour. Drug allergies queried by a yes or no question

and allergic sensitisation evaluated by positive reaction to dust mites and histamine (positive control) showed a strong association to eczema (adjusted odds of 1.45 and 1.72 respectively) indicating an atopic background to the eczema.

2.3.4.3 Evaluating association of family background to eczema

Family background with regards to education and allergic history were also evaluated. Table 2.9 summarises mother's background and history as a covariate for eczema.

Categorical variables	Odds ratio	(95%CI)	p value	Adjusted OR	(95% CI)	p value
Mother's education						
Secondary	Ref	NA	NA	Ref	NA	NA
Primary	0.99	(0.84 - 1.17)	0.947	0.98	(0.83 - 1.16)	0.842
Tertiary	1.11	(0.97 - 1.27)	0.099	1.09	(0.94 - 1.25)	0.227
Mother asthma						
No	Ref	NA	NA	Ref	NA	NA
Yes	1.3	(0.99 - 1.72)	0.056	1.31	(0.98 - 1.75)	0.06
Mother AR						
No	Ref	NA	NA	Ref	NA	NA
Yes	**2.27**	**(1.60 - 3.24)**	**<0.001**	**2.43**	**(1.67 - 3.55)**	**<0.001**
Mother eczema						
No	Ref	NA	NA	Ref	NA	NA
Yes	**2.68**	**(2.20 - 3.27)**	**<0.001**	**2.64**	**(2.14 - 3.24)**	**<0.001**
Mother allergy						
No	Ref	NA	NA	Ref	NA	NA
Yes	**1.94**	**(1.65 - 2.29)**	**<0.001**	**1.89**	**(1.59 - 2.24)**	**<0.001**

Table 2-9: Mother's history evaluated for eczema. Uni-variate and multivariate analyses showing crude and adjusted odds ratio.

Mother's education and asthma status did not show any association to the participant's eczema status. However, mother's allergic rhinitis and eczema were highly associated (adjusted OR 2.4 and 2.6 respectively) to the individual's chances of developing eczema. Although, not a test for risk, the

odds ratio approximate the relative risk making mother's eczema a high risk factor. If the mother had any allergies (Mother allergy) was evaluated by the presence of any of the three queried allergic diseases and it was found to be associated, indicating maternal atopic background predisposes an individual to develop eczema when compared to no maternal atopy.

Similar analyses were done for father's education and allergic history as shown in table 2.10.

Categorical variables	Odds ratio	(95%CI)	p value	Adjusted OR	(95% CI)	p value
Father's education						
Secondary	Ref	NA	NA	Ref	NA	NA
Primary	**1.23**	**(1.04 - 1.47)**	**0.016**	**1.22**	**(1.02 - 1.46)**	**0.024**
Tertiary	1.11	(0.98 - 1.27)	0.097	1.09	(0.96 - 1.25)	0.169
Father asthma						
No	Ref	NA	NA	Ref	NA	NA
Yes	**1.68**	**(1.25 - 2.25)**	**0.001**	**1.65**	**(1.21 - 2.24)**	**0.001**
Father AR						
No	Ref	NA	NA	Ref	NA	NA
Yes	**1.9**	**(1.25 - 2.90)**	**0.003**	**1.97**	**(1.25 - 3.11)**	**0.003**
Father eczema						
No	Ref	NA	NA	Ref	NA	NA
Yes	**3.6**	**(2.86 - 4.53)**	**<0.001**	**3.68**	**(2.91 - 4.67)**	**<0.001**
Father allergy						
No	Ref	NA	NA	Ref	NA	NA
Yes	**2.41**	**(2.01 - 2.88)**	**<0.001**	**2.42**	**(2.01 - 2.91)**	**<0.001**

Table 2-10: Father's history evaluated for eczema. Uni-variate and ultivariate analyses showing crude and adjusted odds ratio.

There was a higher prevalence of eczema in the group whose father's maximum education was until primary as compared to secondary (OR 1.2). Father's asthma, AR status affected the individuals odds of having eczema (OR 1.6 and 1.9 respectively). Paternal eczema was the most highly associated

with the prevalence odds of 3.6 as compared to no paternal eczema. As all allergic diseases were associated with the individual's eczema status, 'father allergy' calculated as presence of either of the three diseases showed association at odds ratio of 2.4.

To check the combined odds for maternal and paternal effect, data was grouped in the same categories as above for disease present in both parents.

The analyses was done with a reference of not having both parents affected by the same allergic disease. Results for the same are tabulated in table 2.11.

Categorical variables	Odds ratio	(95%CI)	p value	Adjusted OR	(95% CI)	p value
Parents asthma						
No	Ref	NA	NA	Ref	NA	NA
Yes	1.41	(0.52 - 3.79)	0.493	1.41	(0.52 - 3.81)	0.488
Parents AR						
No	Ref	NA	NA	Ref	NA	NA
Yes	1.43	(0.47 - 4.31)	0.525	0.93	(0.20 - 4.18)	0.928
Parents eczema						
No	Ref	NA	NA	Ref	NA	NA
Yes	**3.46**	**(2.04 - 5.87)**	**<0.001**	**3.31**	**(1.91 - 5.76)**	**<0.001**
Parents allergy						
No	Ref	NA	NA	Ref	NA	NA
Yes	**2.57**	**(1.84 - 3.59)**	**<0.001**	**2.52**	**(1.77 - 3.60)**	**<0.001**

Table 2-11: Both parent's atopic history evaluated for eczema. Uni-variate and multivariate analyses showing crude and adjusted odds ratio.

Association to eczema was lost when comparing both parents' asthma and AR status. Understandably, parental eczema (OR 3.4) and allergy (OR 2.5) remained significant with an odds ratio between the individual parent's odds.

Information on number of siblings and number of siblings affected with the three allergic diseases were also gathered. Association for number of sibling with allergic diseases to eczema status of the individual was evaluated. A final combined analyses involving presence of allergic disease in all - mother, father and at least one sibling was termed as 'family disease'. Unadjusted and adjusted odds ratio for the same confounding variables consisting of gender, age and race are tabulated in table 2.12 below.

Categorical variables	Odds ratio	(95%CI)	p value	Adjusted OR	(95% CI)	p value
No. of siblings w asthma						
0	Ref	NA	NA	Ref	NA	NA
1	**1.4**	**(1.18 - 1.66)**	**<0.001**	**1.39**	**(1.16 - 1.65)**	**<0.001**
>=2	**1.56**	**(1.16 - 2.09)**	**0.003**	**1.51**	**(1.12 - 2.04)**	**<0.001**
No. of siblings w AR						
0	Ref	NA	NA	Ref	NA	NA
1	**2.1**	**(1.54 - 2.87)**	**<0.001**	**2.3**	**(1.66 - 3.19)**	**<0.001**
>=2	**1.61**	**(1.19 - 2.19)**	**0.002**	**1.51**	**(1.10 - 2.08)**	**0.010**
No. of siblings w eczema						
0	Ref	NA	NA	Ref	NA	NA
1	**2.72**	**(2.26 - 3.28)**	**<0.001**	**2.77**	**(2.29 - 3.36)**	**<0.001**
>=2	**2.75**	**(2.10 - 3.60)**	**<0.001**	**2.48**	**(1.88 - 3.28)**	**<0.001**
Sibling allergy						
No	Ref	NA	NA	Ref	NA	NA
Yes	**1.9**	**(1.67 - 2.16)**	**<0.001**	**1.87**	**(1.64 - 2.14)**	**<0.001**
Family asthma						
No	Ref	NA	NA	Ref	NA	NA
Yes	**1.45**	**(1.27 - 1.67)**	**<0.001**	**1.41**	**(1.23 - 1.62)**	**<0.001**
Family AR						
No	Ref	NA	NA	Ref	NA	NA
Yes	**1.79**	**(1.48 - 2.17)**	**<0.001**	**1.86**	**(1.52 - 2.27)**	**<0.001**
Family eczema						
No	Ref	NA	NA	Ref	NA	NA
Yes	**2.82**	**(2.47 - 3.21)**	**<0.001**	**2.79**	**(2.44 - 3.20)**	**<0.001**
Family allergy						
No	Ref	NA	NA	Ref	NA	NA
Yes	**2.09**	**(1.85 - 2.35)**	**<0.001**	**2.05**	**(1.81 - 2.32)**	**<0.001**

Table 2-12: Sibling and family's atopic history evaluated for eczema. Univariate and multivariate analyses showing crude and adjusted odds ratio.
Abbreviations: No.- number, w - with.

When comparing siblings affected with asthma, the odds of having eczema increased with increase in the number of affected siblings. The same increasing trend is not observed with increasing numbers of siblings affected with AR although they are still significant. Odds are the highest when the siblings also have eczema which doesn't change with change with the number of affected siblings. As all three categories are associated, 'sibling allergy' categorised as presence of at least one of the diseases versus none, is also associated. Family atopic status is also associated with eczema with family eczema being the highest (OR 2.8), followed by family AR (OR 1.8) and then family asthma (OR 1.4).

2.3.4.4 Evaluating association of asthma and AR to eczema

Participant's asthma and AR status were evaluated from the questionnaire. They were classified as a case for asthma if they replied positively to "Have you ever had asthma?" A participant was classified as a case for AR if they had at least two recent symptoms from itchy nose, sneezing, runny or blocked nose without having flu. Analysis was also performed for the presence of both allergic diseases. Results for the same are tabulated in Table 2.13 on the next page.

Both diseases were associated with eczema, with higher prevalence of AR in eczema as compared to asthma. The combined analysis also showed a higher prevalence of eczema in people affected with asthma and AR both as compared to individuals not having both the diseases.

Categorical variables		OR	(95%CI)	p value	Adjusted OR	(95% CI)	p value
Asthma							
	No	Ref	NA	NA	Ref	NA	NA
	Yes	**1.57**	**(1.36 - 1.80)**	**<0.001**	**1.61**	**(1.39 - 1.86)**	**<0.001**
AR							
	No	Ref	NA	NA	Ref	NA	NA
	Yes	**2.03**	**(1.80 - 2.29)**	**<0.001**	**2.06**	**(1.82 - 2.33)**	**<0.001**
Asthma + AR							
	No	Ref	NA	NA	Ref	NA	NA
	Yes	**1.95**	**(1.64 - 2.33)**	**<0.001**	**2.06**	**(1.72 - 2.47)**	**<0.001**

Table 2-13: Asthma and allergic AR evaluated for eczema. Uni-variate and multivariate analyses showing crude and adjusted odds ratio.
Abbreviations: AR- allergic rhinitis.

2.4 Discussion

2.4.1 ISAAC questionnaire

ISAAC was initiated to describe the prevalence and severity of asthma, rhinitis, and eczema in different countries. This was done to obtain baseline measurements in order to be able to assess future trends and form a unanimous basis for future research. The first diagnostic criteria proposed by Hanifin and Rajka (Hanifin and Rajka, 1980) had been widely applied to clinical studies but had not been compatible with epidemiological setting. The UK working party diagnostic criteria started by British dermatologists had started developing and validating diagnosis based on questionnaire data. These questions were incorporated into the core questions of the ISAAC questionnaire (Asher *et al.*, 1995; ISAAC, 1993). Answering chronic itchy rash was used in diagnosis as it helped discriminate the typical mild-moderate atopic dermatitis from non-atopic eczema and other dermatological conditions

with a 91% sensitivity and around 46% specificity. Whereas the question about flexural involvement further enhanced the specificity to 81% and was sensitive at 94% (Williams *et al.*, 1994a; Williams *et al.*, 1994b; Williams *et al.*, 1994c).

2.4.2 Disease definition

According to WAO, 'eczema' is defined as a skin diseases involving a genetically determined skin barrier defect. When atopic constitution of the dermatitis is dominated by IgE antibodies (measured by either IgE levels or skin prick test), then the term or atopic eczema or atopic dermatitis should be used (Johansson *et al.*, 2004). As ISAAC is a measure of eczema and prior studies in Singapore and other countries are also 'eczema' based, we defined our cases for this epidemiological study independent of skin prick results. Thus our disease definition for eczema case is someone who has a chronic itchy flexural rash. This resulting distinction gave us the opportunity to evaluate the association of atopy to eczema.

2.4.3 Population

This epidemiological study is a based on a large cohort of 8334 participants who fitted our age criteria. The male to female ratio was slightly skewed towards the females (56.5%) as the university has a higher population of female students. Although Singapore's ethnic composition is around 74% Chinese, 13% Malays, 9% Indian (www.singstat.gov.sg), we collected a more Chinese heavy cohort with Chinese accounting for 91%, Malays accounting for 2% and 4% Indians. The low contribution of Malays can be justified by the month we conduct our collection. Our collections are usually in the first two-three week of the start of semester in August-September around the time of the

fasting month. As the collection involves mouth-wash which may involve some water being swallowed, Malay students do not readily participate. Other demographic variables such as housing and family income had a good spread across the different categories. Classification based on country of birth showed two- thirds of the cohort to be born in Singapore and about one-third in other countries. This is due to the large international student community in the university and Singapore.

2.4.4 Self-reported prevalence of asthma, rhinitis and eczema

Self reported asthma measures asthma at any stage in one's life, even childhood asthma which one might have outgrown. From our data, asthma (18.45%) correlated very highly with 'ever wheezing' (18.49%), a hallmark of asthma. In 1994, the data presented by an ISAAC study on 6-7 and 12-15 year olds in Singapore showed prevalence of asthma at 18.5% and 20.7% respectively. And a later ISAAC done on the same age group in 2001 showed prevalence of asthma at 16.3% and 27.4% respectively reporting a significant increase in the 12-15 year age groups. However, our data showed the prevalence at 18.7%, 20.3% and 15.5% at the 18-20, 20.01-22 and 22.01-32 age groups. Although a fair comparison is not possible due to the different age groups, it can still safely be implied that the prevalence of asthma has reached a level of stability in Singapore. The oldest age group showed a reduced prevalence of 'ever-asthma' (15.5%), this may be due to a different environment when they were growing up as compared to the environment the 18-22 aged individuals experienced. As compared to the earlier two data sets, wheezing in the last year is much lesser at 5% as compared to 13% (1994) and 10% (2001). This could also be explained by the age difference, as with

increasing age most people grow out of asthma. Within our band of 18-32 years, there did not seem to be much difference as the individuals who grow out of asthma had probably grown out by then keeping the prevalence similar within the three age groups (Goh *et al.*, 1996; Wang *et al.*, 2004).

All symptoms for rhinitis showed a drastic increase when comparing to the two earlier studies. This again could be explained on the age disparity, as rhinitis has a much later onset as compared to asthma. Also, change in environment making people more sensitive, more awareness and better diagnosis for rhinitis could very well be the contributing factors as well. Like in other developed countries, prevalence of atopy related diseases are on the rise due to industrialisation and urbanisation (Goh *et al.*, 1996; Wang *et al.*, 2004).

Chronic rash has shown a steady increase over the years with 10.5% and 12.3% in 1994 to 12.5% and 14.9% in 2001 in 6-7 year olds and 12-15year olds respectively. In our study of 18-32 year olds, the prevalence of chronic rash was found to be 20.98%, a staggering increase considering eczema is a childhood disease. Chronic itchy rash on skin flexures has been showing a steady increase as well, with 6.1% in 1994 to 9.8% in 2001 for the 6-7 year olds and 7% in 1994 to 10.2% in 2001 for the 12-15 year olds and data depicting the numbers as 15.7% for 18-32 year olds. Eczema is an early childhood disease, often said to trigger the atopic march of asthma and allergic rhinitis later on. Like asthma, people are also known to grow out of mild childhood eczema. However, the **increasing numbers** in the 1994 and 2001 study followed by the high prevalence in a much older age group (Goh *et al.*, 1996; Wang *et al.*, 2004) suggests a very strong and active environmental

component to its manifestation. Another study in Singapore schools was done in 2001 using the UK working party's diagnostic criteria which involved a questionnaire and a full body physical examination by one dermatology nurse. The prevalence was found to be 22.7% for 7 year olds, 17.9% for 12 year olds and 21.5% for 16 year olds which the authors argue is a true representation due to the full-body examination (Tay *et al.*, 2002). There has also been a remarkable increase in self reporting of eczema, with 3% and 4% in 1994 to 9% and 6% in 2001. Our data shows this value at 14% for the combined age groups (Goh *et al.*, 1996; Wang *et al.*, 2004). This could be explained by the higher awareness due to age, information advancement and increasing prevalence.

2.4.5 Self-reported prevalence of *severity* for asthma, rhinitis and eczema

Number of wheezing episodes in the past one year have shown a significant decrease in the 6-7 year olds whereas a slight trend of increase has been observed in the 12-15 year olds based on the two prior ISAAC studies. On the other hand, our data shows an increase in prevalence of 4-12 episodes a year (21%) and a similar prevalence of more than 12 episodes a year (5%). The severity of wheezing measured by the number of nights woken up shows a reduced trend as compared to the earlier studies, thus showing reduction in severity with increase in age. This was also highlighted with the drastic decrease in prevalence of speech limitation from a 24% (Goh *et al.*, 1996) and 18% (Wang *et al.*, 2004) in the 12-15 year olds to 12% in the 18-32 year olds.

For rhinitis symptoms, earlier studies have shown a significant higher prevalence of severity in the 12-15 year olds in the 2001 study as compared to

1994 (Goh *et al.*, 1996; Tay *et al.*, 2002). Our data shows a major increase in prevalence of the moderate form of rhinitis affecting daily life as compared to the previous studies. We further found that one fourth find their symptoms troublesome. Also, about 14% of the participants reported the symptoms affected their work and leisure activities and around 22% found the symptoms causing sleep disturbance. Majority of the cohort has been living with the rhinitis conditions for over 10 years (29%) and close to 28% experiencing it in the past five years. This number is largely accounted by the section of cohort who have been in living in Singapore for less than 10 years. These results indirectly show the change in environment affecting a person's allergic status.

Persistent rash was shown to be reducing with age (6-7 year olds versus 12-15 year olds) and with time (1994 versus 2001). The trend seems similar within the three age groups which reduce with age and are lesser than the earlier two studies. The severity of eczema as noted by nights spent awake due to rash increased with age and with the time of the study (Goh *et al.*, 1996; Wang *et al.*, 2004). However, it seemed to have plateau-ed in the higher age group of 18-22 and then tapered down in the last age group of 22.01-32. There was a higher prevalence of late onset eczema. This could be either due to type of collection, as population ascertained in a community setting may have a later onset as compared to hospital based collections (Williams and Strachan, 1998).

2.4.6 Association results

2.4.6.1 Association of demographics to eczema

Females had a higher prevalence of eczema with OR 1.2, although this may be higher or flipped when considering within ethnicities. Higher prevalence in females has been reported in other studies as well (Diepgen and Fartasch, 1992; Goh *et al.*, 1996; Larsen *et al.*, 1996). Interestingly, the older age group of 22.01-32 seemed to have a higher prevalence of eczema. This could again be due to conclusion put forth by Williams and Strachan that a population collection would yield higher number of late onset eczema (Williams and Strachan, 1998). Extending this to the recall bias of these participants to the more recent past as compared to those who had infantile eczema, the numbers would skew towards the older age group. Malays had a higher prevalence with odds of 1.97 as compared to Indians. Eczema was found to be more common in Chinese and Malays as compared to Indians in Singapore (Tay *et al.*, 2002). This was consistent with our results when we controlled for confounding demographics (Table 2.5). Also, studies done in Malaysia, a neighbouring country with an environment similar to Singapore showed the same increased prevalence in Malays and Chinese as compared to Indians (Jaafar and Pettit, 1993). There was no socio-economic effect observed in the prevalence of eczema unlike other allergic diseases. Also, country of birth and years spent in Singapore did not show a difference in prevalence. As demographics such as gender, age and race showed a significant difference, other variables were tested individually and adjusted for the confounding effect of the variables (Table 2.5).

2.4.6.2 Association of diet and lifestyle to eczema

Different foodstuffs showed associations with eczema. Group eating cereals and rice once or twice a week as compared to daily eaters had a higher prevalence of eczema (adjusted OR 1.2 and 1.23 respectively). Eating pulses, nuts and potatoes on most days gave an adjusted OR of 1.17, 1.29 and 1.21 respectively. People who drink yogurt drinks daily had an adjusted OR of 1.2 indicating higher prevalence. Probiotic drinks have been reviewed to have a preventive effect for skin diseases such as eczema, acne, skin hypersensitivity (Ozdemir and Erol, 2013; Roudsari *et al.*, 2013). Amongst the list of foodstuffs evaluated, fast-food was the most highly associated with adjusted OR as 1.37 indicating a high prevalence of eczema in people who consume fast food daily as compared to once or twice a week. As the questionnaire did not ask questions about diet habits before and after onset of the disease due to a possible recall bias, associations of healthy food items as a diet choice cannot be concluded.

Sitting for longer than 5 hours in front of the computer or television or playing video games was associated with higher prevalence of eczema at adjusted OR of 1.29. Unlike other studies (Pelucchi *et al.*, 2013), pet exposure showed no association. Similarly, smoke exposure also showed no association (Morales Suarez-Varela *et al.*, 2008). People with drug allergies reported a significant higher numbers of eczema as compared to people without drug allergies (adjusted OR 1.45) similar to other studies (Baek *et al.*, 2013). As an underlying condition of atopy is a known predisposition factor for developing eczema and/or atopic dermatitis (Chew *et al.*, 1999; Goodwin, 2008; Ronchetti *et al.*, 2003; Stajminger *et al.*, 2007), we found atopy as measured by allergic

sensitisation to common dust mites to be significantly associated with eczema with adjusted OR of 1.72.

2.4.6.3 Association of family background to eczema

While mother's education showed no association, father's primary education group showed a higher prevalence of AD (adjusted OR 1.2). Contrarily, a study in Japan showed father's higher education was a risk factor to eczema (Miyake *et al.*, 2012). As eczema is a genetic disorder, it was obvious to find mother's and father's eczema status strongly associated (Batlles Garrido *et al.*, 2010). However, father's eczema was most strongly associated (adjusted OR 3.68) as compared to mother's eczema (adjusted OR 2.64). Mother's eczema status followed by rhinitis were associated with eczema. There was no association found between maternal asthma and participant's eczema. On the other hand, paternal asthma was associated with eczema although least strongly compared to father's eczema and allergic rhinitis status. Overall an atopic parental background was found to be strongly associated. When comparing presence of the disease in both parents, eczema was the only disease which showed a strong association with adjusted OR of 3.31. Parental allergy defined by the presence of any of the allergic diseases in both the parents was also strongly associated. As with other studies, our results also show a strong genetic link to AD.

The genetic link was further corroborated with similar trends in sibling's disease status. As compared to siblings with no disease, asthma, rhinitis and eczema all showed a positive association. Similarly, family's status as defined by both parents and at least one sibling also showed the strong genetic association to eczema.

2.4.6.4 Association of asthma and rhinitis to eczema

Asthma and allergic rhinitis were also associated with eczema with rhinitis being more strongly associated. Other studies have reported similar associations (Batlles Garrido *et al.*, 2010; Garrett *et al.*, 2013; Silverberg and Hanifin, 2013). Thus the associations with other atopic diseases suggests a set of common underlying candidates for atopy.

2.5 Conclusion

Comparison between the previously reported epidemiology studies and our study showed a steady increase in prevalence of eczema in Singapore. Thus highlighting the effect of changing environment. This study also highlights the genetic link of eczema, with the strongest association found for paternal eczema. Eczema overlaps with asthma and allergic rhinitis in prevalence due to the similar underlying atopic cause. But, like any complex disease, eczema also has environment as a contributing factor besides genetics. Parameters such as gender (female), race (Malay), age, atopy and paternal education seem to be shifting the odds towards developing eczema. Other parameters such as diet and sedentary lifestyle were also shown to be associated.

Chapter 3

Candidate Gene Association Study on Atopic Dermatitis in Singapore Chinese: Screening and Validation

3.1 Introduction

Atopic dermatitis is a polygenic multi-factorial disease. The evidence for genetics in development of any disease classically comes from family studies. Familial aggregation of AD has been long shown from genetic and epidemiological studies since 1982 (Fergusson *et al.*, 1982) to the recent most in 2013 (Martin *et al.*, 2013). Based on previous chapter's results, paternal (OR 3.6) and maternal (OR 3.6) eczema have the highest odds of the child developing eczema, clearly indicating a strong genetic component. Employing ISAAC helped provide a standardised data on prevalence and burden on the population. Thus providing the push needed to reveal the underlying factors for the variations observed in the disease manifestation (Fok and Wong, 2009).

Studies done in Avon county, England show a strong association between parental eczema and childhood AD with an odds ratio of 2.72 (95% CI: 2.09 to 3.53) — both parents affected and odds of 1.7 (95% CI: 1.4 to 2.0) – 1 affected parent (Wadonda-Kabondo, 2004). Twin studies reporting concordance rates of 0.72 - 0.86 in monozygotic and 0.21-0.23 in dizygotic twins further confirmed the genetic contribution to the disease (Schultz Larsen, 1993; Thomsen *et al.*, 2007).

With the reported literature and our epidemiology suggesting genetic factors to be important in the pathogenesis of the disease, then linkage or association studies form the next step to determine the genes responsible. Analysis by linkage identifies broad genomic regions containing the putative disease gene/genes. Linkage analyses relies on co-segregation of genetic loci during meiosis in families. By genotyping regularly spaced out markers on the genome and studying their segregation in affected pedigrees, relative position of the causative regions can be identified (Dawn Teare and Barrett, 2005). In linkage analyses, the causative loci is not measured whereas the spaced-out markers in the genome are measure with respect to the phenotype. Thus the relationship between the marker and the phenotype implies the presence of the causative locus in proximity of the marker. The location of the causative locus relative to the marker is expressed in terms of recombinant fraction which is a non-linear function of the genetic distance between them (Borecki and Province, 2008).

On the other hand, association analyses measures the functional/ causal variant either directly or through another marker which is in linkage disequilibrium (LD) with the functional/ causal variant (Boerwinkle *et al.*, 1986). In association studies, a higher frequency of a single-nucleotide polymorphism (SNP) in the group of individuals affected with a disease can be interpreted as that SNP increases the risk of developing the disease. Microsatellite markers, insertion/deletions, variable-number tandem repeats (VNTRs), and copy-number variants (CNVs) can also be used as the marker in association studies although SNPs are the most widely tested. Broadly speaking, association studies can be classified into four groups, namely, case-control studies in (i)

candidate gene study, (ii) genome wide association studies, (iii) family based association studies and (iv) quantitative trait locus studies (Lewis and Knight, 2012).

Linkage study has many shortcomings, the most important being ascertaining a large number of families with several affected generation. If the disease of interest has a late-onset with a high mortality, finding families with more than one affected generation becomes difficult (Gasser *et al.*, 1998). Generally, linkage studies are less helpful for complex traits where different sets of multiple genes are important in disease causation in different families. Linkage studies are more helpful in identifying loci and genes involved in Mendelian diseases. Also, the resolution of linkage analyses is in centi-Morgans which then further needs deep sequencing and hypothesis for identifying the biologically relevant genes within the loci.

Linkage analysis is effective for relatively low allele frequencies having a strong effect size, whereas association studies cover a larger area of relatively low-high allele frequencies having a moderate to strong effect size. Thus, linkage suits Mendelian diseases better and association studies suit the complex diseases which follow common variant-common disease hypothesis (Ardlie *et al.*, 2002). Besides overcoming the shortcomings of linkage study, association study also allows for investigating gene-gene and gene-environment interactions. Hence, due to the varied advantages of association studies over linkage, in this chapter we focus on the association analysis of AD.

Patterns of LD among SNPs have revealed remarkable distinction between the major populations resulting in the human Haplotype Map (HapMap) project (http://www.hapmap.org/) (Altshuler *et al.*, 2010). Initially, four populations comprising of Han Chinese in Beijing (CHB), Japanese in Tokyo (JPT), Yoruba in Ibadan, Nigeria (YRI) and Utah residents with ancestry from northern and western Europe (CEU) were studied and their detailed allele frequencies and LD pattern made available. The LD pattern has been shown to be stable across different individuals aiding in analysing the genome. The resulting correlated alleles within haplotype blocks identifies tag SNPs (t-SNPs) which form a set of fewer reliable markers within a population for analysing their entire genome. This understanding of the structure within the human genome has made the genome wide studies on a panel/chip possible. This information is also helpful in making one's own customised panel for a candidate gene study (Borecki and Province, 2008) and are best suited for identifying differences for common diseases (Kruglyak and Nickerson, 2001; Nakamura, 2009).

Based on the justifications and the ease of conducting a candidate gene study with a prior well characterised epidemiology population ascertained, I conducted a candidate gene screening study on the ethnic Chinese population of Singapore. The genes were selected based on the barrier dysfunction hypothesis. Genes were identified either through literature or through biological plausibility of the gene's product to have a role in maintaining the skin's barrier. A comprehensive list of 222 genes was generated. However, due to technical and financial limitations only 145 of them could be genotyped. Weakly associated, and genes not well replicated in other studies were not

selected in the reduced 145. The selection for the 145 list was also based on the subjective importance given to the gene in pathogenesis of AD. Most of the genes worked upon were from chromosome 1 which contains the epidermal differentiation complex (Cookson *et al.*, 2001) and filaggrin- the most widely associated gene to AD (Barker *et al.*, 2007; Palmer *et al.*, 2006; Weidinger *et al.*, 2006)

Although genome wide association (GWAS) techniques are widely used, we did not adopt it, as GWAS needs large sample sizes to conduct a well powered study. Ascertaining around 1000 cases would have meant ascertaining 10,000 samples as AD was found to be prevalent around 10% in our population. Also, as my study was hypothesis-driven, unlike a GWAS, we adopted a candidate gene approach. We, however, validated the Chinese GWAS in our population resulting in a paper as described in Chapter 5. Studies done in our lab (Andiappan *et al.*, 2010) have shown than the ethnic Chinese population ascertained by us is 95% correlated to the HapMap's CHB population. Hence, the CHB database was used as the reference population for identifying SNPs in this candidate gene study.

3.2 Materials and methods

3.2.1 Population ascertainment

Population was ascertained using a cross-sectional sampling methodology from National University of Singapore's (NUS) campus and a paediatric allergy clinic from KK Women's and Children's Hospital (KKH). Cases and controls for AD were identified and selected based on the questionnaire and

skin prick test's (SPT) results. In addition to the samples collected from NUS for the epidemiological study, samples were also recruited from an allergy clinic at KKH. Approval to conduct this recruitment at KKH was obtained from the Singhealth Centralised Institutional Review Board. Parental or guardian consent was obtained for participants below the age of 21 years and $10 vouchers were given as a reimbursement for their participation.

3.2.2 Disease definition

For disease classification, participants filled out a standardised and validated International Study of Asthma and Allergies in Childhood (ISAAC) questionnaire (Asher *et al.*, 1995) and underwent a skin prick test (SPT). Categorisation into the disease bin for either of the three allergic diseases required a positive SPT to common allergens found in Singapore. The participants were tested for four aeroallergens namely *Dermatophagoides pteronyssinus, Blomia tropicalis, Elaeis guineensis* and *Curvularia lunata*. To validate the test a positive and negative control, i.e., histamine and saline (buffer used for the preparation of the allergen) respectively, were also skin pricked. Studies done by the lab previously showed that skin-prick reactivity to dust mites especially the two species *Blomia tropicalis* and *Dermatophagoides pteronyssinus* were highly sensitive and specific markers for allergic sensitization in Singapore (Chew *et al.*, 1999). SPT was considered positive when the wheal diameter was ≥ 3mm, whereas no signs of wheal and erythema was considered a negative SPT. To ensure the SPT is being carried out correctly, a positive reaction to histamine and a negative reaction to saline was mandatory before the sample being considered as either a case or a control of any of the diseases. Together with the positive and negative

controls, a positive reaction to either of the dust mites indicated a positive SPT for the participant. Participants with less than 3mm reaction for histamine or, erythema or wheal to saline were not considered for defining the cases and controls. Participants with smaller than 3mm reaction to the test allergens were considered to have an intermediate SPT result.

A participant was categorised as a case for AD if he/she answered affirmative to "ever had itchy rash" and "rash at skin flexures" in addition to testing positive for SPT. Based on UK working party's criteria, flexural involvement of rashes had 90% sensitivity and 75% specificity for diagnosing AD (Williams *et al.*, 1994a). A subject was categorized as a 'control' for AD if he/she answered negative to the above questions in addition to a negative response to "ever had eczema" with a negative SPT. A participant was classified as a case for asthma if they replied positively to "Have you ever had asthma?" and had a positive skin prick result. A participant was classified as a case for AR if they had at least two recent symptoms from itchy nose, sneezing, runny or blocked nose without having flu and a positive SPT result.

In addition to case and control, 'SPT positive' were samples which had a positive reaction to the skin prick test, however did not have disease according to the questionnaire. 'Phenotype unknown' were samples who could not be classified either as a case or control due to their responses to the questionnaire. 'SPT inconclusive' were samples which had inconclusive skin prick test results and 'Saline reaction' were samples which reacted positively to the negative control.

3.2.3 Population

For the screening study, 1090 samples comprising of 449 AD cases and 641 controls were genotyped. The significantly associated genes were then validated on a larger population of 732 AD cases, 1085 controls and 650 SPT positives & survey negative for AD.

Samples collected from NUS and KKH were further stratified according to severity and chronicity. Severity was defined by the number of nights they were kept awake by the itchy rash; mild cases were those that were never disturbed at night, moderate were those that were kept awake less than one night per week in the last one year and severe were those which cases that were kept awake for more than one night per week.

Chronicity was defined based on the participant's response to two questions- namely, presence of rash in the last one year and complete clearance of the rash in this past year. Chronic cases were identified as those who replied 'Yes' to rash and 'No' to its complete clearing. Similarly, non-chronic cases were those who replied 'Yes' to complete clearance when they had a rash in the past year, and those who did not have a rash in the past one year . Samples which did not provide answers in this format or had missing information to these two questions were excluded.

3.2.4 DNA collection from participants

Mouth-wash sample was collected from participants recruited from the NUS campus. Buccal cells collected in 10ml of 0.9% saline were pelleted by centrifuging at 5,000 rpm at 4°C for 15mins. The supernatant was discarded and the cells were re-suspended in 500µl of digestion buffer (100mM NaCl,

10mM TrisCl at pH8, 25mM EDTA at pH8, 0.5% SDS) and subsequently incubated at 37°C overnight in a water bath. Following which steps for genomic DNA extraction were performed according to the manual provided with AxyPrep Multisource Genomic DNA Miniprep Kit (Axygen). Eluted DNA was then quantified in duplicate on the Nanodrop ND-2000 (Thermo Scientific). Only samples with DNA absorbance ratios greater than 1.8 at 260nm/280nm, greater than 1.5 at 260nm/230nm, and with concentration higher than 20ng/μl were considered to be of good quality and used for downstream genetic analyses. KKH participants provided DNA via a blood sample as obtaining a mouth wash samples from the paediatric patients was difficult. Hospital's phlebotomist collected a maximum of 2ml blood in sodium-EDTA tubes. DNA was extracted according to the manufacturer's instructions of the Wizard Genomic DNA Purification Kit (Promega). The DNA samples were similarly analyzed and quantified as mentioned above.

3.2.5 Whole Genome Amplification

Whole genome amplification was performed on all cases, controls, SPT positive samples and the severe AD cases from NSC. We used Illustra GenomiPhi HY DNA Amplification Kit (GE Healthcare) to amplify the samples. This kit is based on the multiple displacement amplification principle (Dean *et al.*, 2002), which allows for the uniform amplification of the whole genome producing high yields of DNA amenable to standard genetic analyses techniques such as PCR, sequencing and microarrays (Hosono *et al.*, 2003). The amplified DNA was also quantified and analysed as mentioned above. Whole-genome amplified DNA samples were used for the candidate gene study while the stock DNA samples were used for the methylation study.

3.2.6 Selection of SNPs for Genotyping

Tag SNPs were identified from the genes of interest spanning the gene, 2 kb upstream and 1 kb downstream. Tag SNPs were selected using Tagger (de Bakker *et al.*, 2005) software integrated with HapMap. Tagging was done based on LD. Tag SNPs' selection criteria was a minimum allele frequency (MAF) of 0.05 and LD r^2 of 0.8 within the CHB population. At the time of study, the available full data set release #24 of HapMap was used. The 145 genes selected were tagged completely by 978 SNPs.

3.2.7 Genotyping on the Illumina BeadXpress

Genotyping was outsourced to the DNA Sequencing and Genomics Facility at the University of Utah and University of Washington for genotyping on the BeadXpress platform using the GoldenGate assay (Illumina)

3.2.8 Statistical Analyses

Data and statistical analyses were performed in Microsoft Excel (http://office.microsoft.com/en-us/excel/) and online calculators (http://www.physics.csbsju.edu/cgi-bin/stats/contingency). Genotyping data was analysed using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/) (Purcell *et al.*, 2007) and R (http://www.R-project.org). The model association test in PLINK tested for general, multiplicative, additive, dominant and recessive models of association (Lewis, 2002). The general model tests the differences between genotype counts, the multiplicative model tests for differences on allelic counts, both using Fisher's exact test. The additive model uses the Cochran Armitage Trend test on allele counts whereas the dominant and recessive models tests for dominant and recessive effect of the minor allele via Fisher's exact test. Association was tested for all models as no prior

assumptions were made. However, only the general and multiplicative models were considered for determining whether an association was present or not. SNP call rates and Hardy-Weinberg Equilibrium were used as quality control measures. SNPs with a Hardy Weinberg Equilibrium *P* value of less than $1x10^{-3}$ were excluded as they were likely to have genotyping errors.

Based on the previous chapter on epidemiology of AD in Singapore, demographics such as gender, age group and race were shown to be associated with the disease and as such may become confounding variables in this genetic study. As this study was Chinese race specific, confounding effect of gender and age group were tested to check for true association of the SNP to disease. This was tested using SPSS (IBM software).

3.3 Results

3.3.1 Screening study

3.3.1.1 Population

Atopic dermatitis is a low prevalent disease as compared to the other allergic disease such as asthma and allergic rhinitis. To establish a sizeable cohort for the screening study, samples from the campus recruitment along with the paediatric allergy clinic of KKH were ascertained resulting in 449 cases and 641 controls. At the time of the screening study, we had collected 5154 samples of which 4604 (89.33%) were ethnic Chinese. Based on the DNA quality and exclusion criteria, as described in the materials and method, 4550 samples had good quality DNA (1.17% loss due to bad quality and low yield).

The classification of the population is as shown in the following Venn diagram (Fig 3.1)

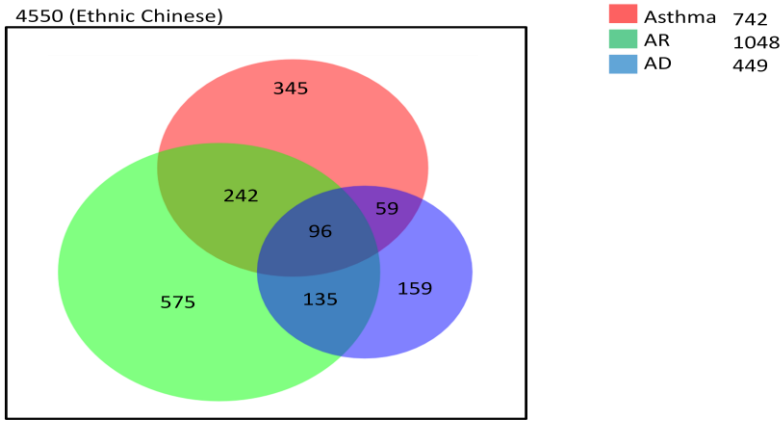

Figure 3-1: Population ascertained for the screening study.
Venn diagram showing the overlapping populations of asthma, allergic rhinitis (AR) and AD. The AD population (449) was used for the screening study.

By ascertaining 4550 good quality DNA from ethnic Chinese we collected 742 (16.3%) asthma cases, 1048 (23%) allergic rhinitis cases and a mere 449 (9.8%) AD cases. Further stratification of the population with respect to AD is as shown below. The definition of the classifying terms is as explained in the materials and method.

	Atopic Dermatitis	
	#	%
Cases	449	9.87
Controls	641	14.09
SPT positive	2047	44.99
Phenotype unknown	574	12.62
SPT Inconclusive	577	12.68
Indeterminate	129	2.84
Saline Reaction	133	2.92
Total	4550	100.00

Table 3-1: Ethnic Chinese population's stratification for screening study based on the AD questionnaire.

From the population of 4550, 641 (14%) samples fitted our definition of AD controls whereas around 45% of the population was skin prick positive to the dust mites but responded negatively to the AD questions indicating a very high sensitisation of the ethnic Chinese to dust mites. A sizeable proportion of 12% replied ambiguously to the AD questions thus rendering them as Phenotype unknown. A similar proportion of the population had inconclusive results to the SPT i.e., developing a wheal of less than 3 mm for histamine and/or the allergens. Close to 3% of the population were classified as AD indeterminate as they replied positively to the AD questions, but were SPT negative and a similar proportion reacted positively to saline hence were excluded from the case and control bins.

As the study population was ethnic Chinese, we relaxed the age groups for the collection. The following table describes the screening population of 449 AD cases and 641 controls.

		Case		Control	
		Number	(%)	Number	(%)
Gender	Male	195	(43.43)	177	(27.61)
	Female	254	(56.57)	464	(72.39)
Age group	< 18	24	(5.35)	3	(0.47)
	18 - 20	146	(32.52)	255	(39.78)
	20.01 - 22	136	(30.29)	191	(29.80)
	22.01 - 32	138	(30.73)	176	(27.46)
	> 32	5	(1.11)	16	(2.50)
Total		449	(100)	641	(100)

Table 3-2: Demographics of the screening population.
AD cases and controls stratified according to gender and age group for the screening study.

The sex ratio in NUS is skewed towards females which is reflected in our recruitment. This has resulted in a higher proportion of females in cases and controls with a staggering 72% females accounting for the controls. The age band from 18-32 have similar numbers in cases and controls. However, there is a higher proportion of cases in less than 18 age group which can be explained by the KKH recruitment. The KKH recruitment was conducted in the allergy clinic of the paediatric ward resulting in the skew. Senior graduate students and staff of NUS accounted for the greater than 32 age group.

3.3.1.2 Genes and tag SNPs

A comprehensive list of genes was identified to be tested for association with AD. The selection was based on (i) genes reported to be associated with AD, (ii) genes coding for proteins associated with AD, (iii) genes associated with other barrier dysfunctional diseases, and lastly (iv) selective gene coding proteins which were discovered by a senior PhD student in the sweat proteome (Manesh 2009, PhD thesis). For the sweat proteome study, sweat-dripping method was the method of sample collection, in which sweat which runs over skin and drips was used for further proteomic studies using 2DE and 2D-LC-MS/MS. As skin run-off sweat was collected, it also contained soluble proteins of the skin. On careful evaluation of the proteins and their known or putative functions 61 genes were identified. This list overlapped with genes/ proteins matching the first three criteria for gene selection. Hence, all the reported novel associations were obtained from the sweat proteome which had no prior reference or context to the skin or AD or barrier maintenance in literature.

A total of 222 genes were identified. Due to limitation of number of SNPs that could be probed and financial constraints, we decided to proceed with genotyping 145 genes of the 222 genes. The 145 genes were prioritised based on the their impact on barrier dysfunction and well documented/ hypothesized candidates. Tagging was adopted to reduced the number of SNPs typed. Tagging was performed based on r^2 (LD) of 0.8 and for SNPs with minor allele frequency (MAF) of 0.05 and above. The genes were tagged 2 kb upstream and 1 kb downstream resulting in 978 tag SNPs representing 145 genes as shown in Tables 3.3 & 3.4. As evident, chromosome 1 housing the epidermal differentiation complex accounts for the maximum number of genes within the list.

Chr	Tags	Count of genes	Chr	Tags	Count of gene
1	117	27	11	113	12
2	18	3	12	40	10
3	93	7	14	45	7
4	8	2	15	22	2
5	47	5	16	15	2
6	125	10	17	25	9
7	4	2	18	120	12
8	52	10	19	57	10
9	20	4	20	36	7
10	17	3	21	4	1
			Total	978	145

Table 3-3: Genes genotyped per chromosome for screening.
Chromosome (Chr) wise stratification of the 145 genes represented by 978 tag SNPs that were genotyped in the screening study.

Table 3-4: Genes and tag SNPs genotyped in the screening study.

Chr	Gene	Tag SNPs	Chr	Gene	Tag SNPs	Chr	Gene	Tag SNPs	Chr	Gene	Tag SNPs	Chr	Gene	Tag SNPs
1	*CA6*	24	2	*ZAP70*	12	8	*CLDN23*	1	12	*KRT2*	5	18	*DSC2*	9
1	*CAP1*	13	3	*CD80*	23	8	*CLU*	6	12	*KRT4*	6	18	*DSC3*	27
1	*CHI3L1*	13	3	*CD86*	13	8	*CTSB*	6	12	*KRT5*	5	18	*DSG1*	17
1	*CHI3L2*	5	3	*CLDN1*	16	8	*DEFA5*	3	12	*KRT6A*	9	18	*DSG2*	16
1	*CREG1*	4	3	*CSTA*	3	8	*DEFA6*	3	12	*KRT6B*	5	18	*DSG3*	12
1	*CRNN*	2	3	*LTF*	23	8	*DEFB104B*	1	12	*LYZ*	1	18	*DSG4*	5
1	*CTBS*	3	3	*TF*	14	8	*LRRC14*	2	12	*MUCL1*	2	18	*SERPINB10*	4
1	*FLG*	6	3	*TLR9*	1	8	*PLEC1*	8	12	*PEBP1*	2	18	*SERPINB12*	4
1	*FLG2*	3	4	*FABP2*	6	9	*CTSL2*	2	14	*CALM1*	5	18	*SERPINB3*	2
1	*HRNR*	4	4	*TLR2*	2	9	*LCN1*	2	14	*CMA1*	7	18	*SERPINB4*	2
1	*IVL*	1	5	*CD14*	1	9	*ORM2*	2	14	*RNASE2*	3	18	*SERPINB7*	13
1	*LOR*	2	5	*GM2A*	11	9	*TXN*	14	14	*RNASE3*	4	19	*CALM3*	3
1	*PDZK1*	4	5	*HEXB*	9	10	*ASAH2*	13	14	*RNASE7*	5	19	*CASP14*	2
1	*PDZK1IP1*	3	5	*IL7R*	11	10	*CALML5*	2	14	*SERPINA12*	20	19	*KLK10*	11
1	*PGLYRP3*	4	5	*SPINK5*	15	10	*VIM*	2	14	*SGPP1*	1	19	*KLK11*	5
1	*RPTN*	1	6	*CDSN*	18	11	*C11orf30*	9	15	*ANXA2*	15	19	*KLK13*	9
1	*S100A11*	2	6	*DSP*	23	11	*CAPN1*	8	15	*CTSH*	7	19	*KLK14*	7
1	*S100A2*	1	6	*OPRM1*	41	11	*CD44*	45	16	*PPL*	13	19	*KLK5*	11
1	*S100A7*	3	6	*PSMB8*	2	11	*CST6*	1	16	*PRSS8*	2	19	*KLK6*	3
1	*S100A7A*	2	6	*PSMB9*	3	11	*CTSC*	14	17	*EVPL*	3	19	*KLK7*	2
1	*S100A8*	1	6	*SERPINB6*	15	11	*CTSD*	3	17	*GRN*	8	19	*KLK8*	4
1	*S100A9*	8	6	*SMPD2*	2	11	*FADS1*	2	17	*KRT10*	2	20	*CST1*	1
1	*SPRR1A*	1	6	*SMPDL3A*	6	11	*FADS2*	8	17	*KRT14*	2	20	*CST4*	2
1	*SPRR1B*	1	6	*VNN1*	12	11	*LRRC32*	6	17	*KRT16*	2	20	*CTSA*	5
1	*SPRR2C*	2	6	*VNN3*	3	11	*MMP10*	5	17	*KRT17*	1	20	*PI3*	2
1	*SPRR2G*	3	7	*AZGP1*	1	11	*PRCP*	11	17	*KRT9*	2	20	*SLPI*	2
1	*TCHH*	1	7	*PIP*	3	11	*SMPD1*	1	17	*SOCS3*	3	20	*TGM3*	20
2	*CALM2*	5	8	*ASAH1*	20	12	*DCD*	2	17	*SPHK1*	2	20	*ZGPAT*	4
2	*DES*	1	8	*CA2*	2	12	*KRT1*	3	18	*DSC1*	9	21	*SOD1*	4

3.3.1.3 Genotyping results

Analyses for all models of genetic association were tested on the 978 SNPs. SNPs associated at a *P* value of less than 0.05 at either the general or multiplicative models were considered significant. A brief summary of the chromosome-wise division of the significantly associated tags and genes is as shown in table 3.5. Table 3.6 shows only the associated 97 tag SNPs from the 46 significant genes.

Chr	Sig tags	Count of sig genes
1	19	5
2	4	2
3	6	3
5	4	2
6	16	5
8	5	2
9	4	2
10	1	1
11	8	4
12	5	3
14	4	3
15	5	2
16	2	1
17	1	1
18	5	4
19	6	4
20	2	2
Total	97	46

Table 3-5: Summary of the significant genes and SNPs per chromosome.
Summary of the 46 genes and the number of tag SNPs found significant (sig) at the screening study

Chr	SNP	Gene	Case	Control	General	aa vs aA		aa vs AA		aA vs AA		Multiplicative model		
					P value	OR	95% CI	OR	95% CI	OR	95% CI	P value	OR	95% CI
1	rs7545200	*CA6*	38/172/231	44/212/383	**0.046**	1.064	(0.66-1.718)	1.432	(0.901-2.277)	**1.345**	**(1.038-1.743)**	**0.015**	**1.275**	**(1.049-1.551)**
1	rs17032907	*CA6*	56/236/151	109/292/237	**0.029**	**0.636**	**(0.441-0.916)**	0.806	(0.551-1.181)	1.269	(0.972-1.656)	0.747	0.972	(0.815-1.158)
1	rs11576766	*CA6*	39/223/184	67/261/311	**0.012**	0.681	(0.442-1.051)	0.984	(0.637-1.52)	**1.444**	**(1.119-1.864)**	0.164	1.139	(0.949-1.367)
1	rs942969	*CA6*	6/113/327	17/192/431	0.064	0.600	(0.23-1.565)	0.465	(0.181-1.193)	0.776	(0.59-1.019)	**0.023**	**0.760**	**(0.599-0.964)**
1	rs3765965	*CA6*	15/120/309	8/144/488	**0.009**	2.250	(0.922-5.488)	**2.961**	**(1.241-7.067)**	1.316	(0.994-1.743)	**0.004**	**1.423**	**(1.118-1.811)**
1	rs10864376	*CA6*	73/194/176	76/281/282	0.075	1.391	(0.961-2.014)	1.539	(1.061-2.233)	1.106	(0.851-1.438)	**0.032**	**1.215**	**(1.017-1.452)**
1	rs2274327	*CA6*	31/186/224	38/227/373	**0.045**	0.996	(0.596-1.662)	1.358	(0.822-2.245)	**1.364**	**(1.057-1.761)**	**0.022**	**1.256**	**(1.033-1.527)**
1	rs2274328	*CA6*	81/241/124	108/305/225	**0.033**	0.949	(0.68-1.326)	1.361	(0.948-1.955)	**1.434**	**(1.087-1.891)**	**0.044**	**1.194**	**(1.005-1.42)**
1	rs2274333	*CA6*	105/225/118	122/319/199	0.107	1.220	(0.893-1.667)	1.451	(1.027-2.052)	1.189	(0.895-1.581)	**0.035**	**1.202**	**(1.013-1.426)**
1	rs12738365	*CA6*	25/192/230	36/227/376	**0.042**	0.821	(0.476-1.416)	1.135	(0.664-1.94)	**1.383**	**(1.074-1.78)**	0.051	1.215	(0.999-1.479)
1	rs17032921	*CA6*	18/127/299	13/158/469	**0.037**	1.723	(0.813-3.649)	**2.172**	**(1.049-4.498)**	1.261	(0.958-1.66)	**0.013**	**1.339**	**(1.063-1.687)**
1	rs16826775	*CAP1*	6/129/309	18/140/482	**0.010**	**0.362**	**(0.139-0.94)**	0.520	(0.204-1.324)	**1.437**	**(1.088-1.899)**	0.168	1.184	(0.931-1.505)
1	rs4284253	*CAP1*	6/128/308	17/141/480	**0.017**	0.389	(0.149-1.016)	0.550	(0.215-1.41)	**1.415**	**(1.071-1.87)**	0.169	1.184	(0.93-1.506)
1	rs7515776	*CHI3L1*	30/203/213	53/235/348	**0.018**	0.655	(0.403-1.065)	0.925	(0.573-1.493)	**1.411**	**(1.095-1.819)**	0.172	1.142	(0.944-1.38)
1	rs946261	*CHI3L1*	44/197/202	93/281/263	0.060	0.675	(0.451-1.009)	0.616	(0.412-0.922)	0.913	(0.705-1.182)	**0.031**	**0.819**	**(0.684-0.982)**
1	rs7542294	*CHI3L1*	77/232/132	91/295/247	**0.008**	1.076	(0.759-1.525)	**1.583**	**(1.094-2.292)**	**1.472**	**(1.121-1.932)**	**0.005**	**1.287**	**(1.081-1.533)**
1	rs2282302	*FLG2*	17/119/307	9/182/448	**0.035**	**2.889**	**(1.247-6.694)**	**2.756**	**(1.213-6.264)**	0.954	(0.726-1.254)	0.316	1.125	(0.893-1.417)
1	rs5509194	*SPRR2G*	61/229/158	61/315/264	**0.038**	1.376	(0.928-2.039)	**1.671**	**(1.113-2.508)**	1.215	(0.936-1.576)	**0.016**	**1.242**	**(1.041-1.483)**
1	rs513555	*SPRR2G*	61/214/169	66/288/286	0.051	1.244	(0.842-1.838)	1.564	(1.052-2.326)	1.257	(0.97-1.63)	**0.016**	**1.246**	**(1.042-1.491)**
2	rs1723484	*CALM2*	5/119/324	4/126/511	**0.015**	1.324	(0.347-5.047)	1.971	(0.526-7.396)	1.490	(1.118-1.984)	**0.005**	**1.441**	**(1.113-1.866)**
2	rs1027478	*CALM2*	38/202/204	35/252/351	**0.006**	1.354	(0.826-2.222)	**1.868**	**(1.144-3.051)**	**1.379**	**(1.071-1.776)**	**0.002**	**1.350**	**(1.117-1.633)**
2	rs1693869	*CALM2*	3/86/359	1/87/551	**0.012**	3.035	(0.31-29.753)	4.604	(0.477-44.44)	1.517	(1.095-2.103)	**0.006**	**1.529**	**(1.127-2.074)**
2	rs1020396	*ZAP70*	22/194/229	51/233/355	**0.020**	**0.518**	**(0.303-0.885)**	0.669	(0.395-1.133)	**1.291**	**(1.002-1.662)**	0.784	1.028	(0.846-1.247)
3	rs6810204	*CD80*	50/215/180	58/284/297	0.120	1.139	(0.75-1.729)	1.422	(0.934-2.167)	1.249	(0.967-1.614)	**0.046**	**1.202**	**(1.003-1.441)**
3	rs6778945	*CD80*	18/100/329	10/148/481	**0.042**	**2.664**	**(1.181-6.01)**	**2.632**	**(1.2-5.773)**	0.988	(0.739-1.32)	0.172	1.185	(0.929-1.513)
3	rs3792287	*CD86*	2/74/370	7/140/491	**0.043**	0.541	(0.11-2.668)	0.379	(0.078-1.836)	0.701	(0.513-0.959)	**0.014**	**0.698**	**(0.524-0.93)**
3	rs6776378	*CLDN1*	9/101/335	31/188/420	**0.001**	0.540	(0.248-1.18)	**0.364**	**(0.171-0.775)**	**0.674**	**(0.508-0.892)**	**0.000**	**0.635**	**(0.501-0.805)**
3	rs3774032	*CLDN1*	1/44/397	2/103/535	**0.006**	1.170	(0.103-13.246)	0.674	(0.061-7.457)	0.576	(0.395-0.839)	**0.005**	**0.602**	**(0.421-0.86)**
3	rs9839711	*CLDN1*	3/67/377	5/130/504	0.071	1.164	(0.27-5.02)	0.802	(0.191-3.377)	0.689	(0.499-0.952)	**0.031**	**0.723**	**(0.537-0.973)**
5	rs6861766	*GM2A*	28/192/226	49/226/365	**0.034**	0.673	(0.407-1.112)	0.923	(0.564-1.511)	**1.372**	**(1.064-1.769)**	0.195	1.136	(0.937-1.378)
5	rs1494560	*IL7R*	54/239/155	106/343/192	0.067	0.731	(0.507-1.055)	0.631	(0.427-0.932)	0.863	(0.66-1.129)	**0.033**	**0.828**	**(0.696-0.985)**
5	rs10213865	*IL7R*	19/145/285	43/220/373	0.112	0.670	(0.376-1.196)	0.578	(0.33-1.014)	0.863	(0.665-1.119)	**0.043**	**0.808**	**(0.657-0.994)**
5	rs6893142	*IL7R*	75/230/137	145/316/178	0.066	0.711	(0.513-0.985)	0.672	(0.47-0.96)	0.946	(0.715-1.251)	**0.042**	**0.836**	**(0.704-0.994)**

Table 3-6: Analyses for the significant tag SNPs at the screening study.

					General	aa vs aA		aa vs AA		aA vs AA		Multiplicative model		
Chr	SNP	Gene	Case	Control	P value	OR	95% CI	OR	95% CI	OR	95% CI	P value	OR	95% CI
6	rs12207756	*CDSN*	0/2/439	0/14/626	**0.021**	NA		NA		0.204	(0.046-0.901)	**0.021**	**0.206**	**(0.047-0.907)**
6	rs2295859	*DSP*	3/74/370	5/73/560	0.055	0.592	(0.136-2.568)	0.908	(0.216-3.823)	1.534	(1.082-2.175)	**0.034**	**1.413**	**(1.026-1.945)**
6	rs2299032	*DSP*	58/207/177	59/293/282	0.095	1.391	(0.929-2.084)	1.566	(1.041-2.356)	1.126	(0.869-1.458)	**0.047**	**1.201**	**(1.002-1.438)**
6	rs926411	*DSP*	14/180/254	27/208/405	**0.029**	0.599	(0.305-1.178)	0.827	(0.425-1.607)	**1.380**	**(1.071-1.779)**	0.126	1.175	(0.956-1.444)
6	rs7755659	*OPRM1*	44/181/222	52/307/279	**0.044**	1.435	(0.923-2.232)	1.063	(0.686-1.649)	**0.741**	**(0.574-0.956)**	0.295	0.906	(0.753-1.09)
6	rs9479769	*OPRM1*	31/208/207	73/265/299	**0.029**	**0.541**	**(0.342-0.855)**	**0.613**	**(0.389-0.968)**	1.134	(0.88-1.461)	0.326	0.911	(0.758-1.097)
6	rs790259	*OPRM1*	36/224/184	67/262/307	**0.010**	**0.628**	**(0.404-0.978)**	0.896	(0.575-1.398)	**1.426**	**(1.105-1.841)**	0.281	1.106	(0.921-1.328)
6	rs4959749	*SERPINB6*	82/236/130	153/357/131	**0.002**	0.811	(0.592-1.111)	**0.540**	**(0.376-0.775)**	**0.666**	**(0.497-0.893)**	**0.001**	**0.753**	**(0.634-0.894)**
6	rs1125455	*SERPINB6*	69/216/161	128/342/167	**0.002**	0.854	(0.608-1.198)	**0.559**	**(0.389-0.805)**	**0.655**	**(0.497-0.863)**	**0.001**	**0.744**	**(0.625-0.885)**
6	rs9503348	*SERPINB6*	79/219/143	142/339/155	**0.010**	0.861	(0.623-1.19)	**0.603**	**(0.422-0.862)**	**0.700**	**(0.527-0.93)**	**0.004**	**0.778**	**(0.654-0.924)**
6	rs2580109	*SERPINB6*	20/169/260	19/193/428	**0.009**	1.202	(0.621-2.328)	1.733	(0.908-3.308)	**1.441**	**(1.114-1.865)**	**0.003**	**1.377**	**(1.116-1.7)**
6	rs2165123	*SERPINB6*	20/170/259	20/198/423	**0.018**	1.165	(0.606-2.237)	1.633	(0.862-3.094)	**1.402**	**(1.085-1.813)**	**0.006**	**1.339**	**(1.086-1.65)**
6	rs7743193	*SERPINB6*	33/206/202	47/246/340	**0.030**	0.838	(0.518-1.358)	1.182	(0.733-1.906)	**1.409**	**(1.093-1.817)**	**0.044**	**1.214**	**(1.005-1.468)**
6	rs2244693	*SERPINB6*	20/166/257	24/193/418	**0.033**	0.969	(0.517-1.817)	1.355	(0.734-2.503)	**1.399**	**(1.08-1.813)**	**0.016**	**1.293**	**(1.049-1.595)**
6	rs2295767	*SERPINB6*	19/151/277	17/173/447	**0.014**	1.280	(0.642-2.552)	1.804	(0.922-3.529)	**1.409**	**(1.081-1.835)**	**0.004**	**1.382**	**(1.11-1.72)**
6	rs13192569	*SMPDL3A*	8/140/296	21/156/459	**0.018**	**0.424**	**(0.182-0.989)**	0.591	(0.258-1.351)	**1.392**	**(1.061-1.825)**	0.216	1.156	(0.919-1.455)
8	rs7508	*ASAH1*	77/241/130	133/285/221	**0.012**	**0.685**	**(0.493-0.951)**	0.984	(0.69-1.403)	**1.438**	**(1.09-1.895)**	0.653	1.040	(0.876-1.236)
8	rs11986226	*ASAH1*	28/154/262	47/259/332	0.077	1.002	(0.602-1.666)	0.755	(0.46-1.239)	0.753	(0.583-0.974)	**0.036**	**0.810**	**(0.665-0.987)**
8	rs7464572	*PLEC1*	12/129/306	10/155/475	0.081	1.442	(0.603-3.445)	1.863	(0.795-4.364)	1.292	(0.982-1.7)	**0.027**	**1.304**	**(1.03-1.651)**
8	rs7010330	*PLEC1*	17/125/306	13/148/477	**0.034**	1.548	(0.724-3.312)	2.038	(0.976-4.257)	1.317	(0.997-1.739)	**0.009**	**1.366**	**(1.081-1.727)**
8	rs11783655	*PLEC1*	16/137/294	14/164/463	0.053	1.368	(0.645-2.903)	1.800	(0.866-3.742)	1.316	(1.004-1.723)	**0.015**	**1.323**	**(1.055-1.661)**
9	rs13284515	*LCN1*	6/101/339	7/113/518	0.119	0.959	(0.312-2.948)	1.310	(0.436-3.931)	1.366	(1.01-1.846)	**0.047**	**1.312**	**(1.003-1.718)**
9	rs4135168	*TXN*	3/40/402	0/89/550	**0.004**	NA		NA		0.615	(0.414-0.912)	0.089	0.728	(0.505-1.051)
9	rs4135188	*TXN*	72/189/178	94/327/216	**0.026**	1.325	(0.929-1.891)	0.929	(0.645-1.34)	**0.701**	**(0.537-0.916)**	0.244	0.900	(0.755-1.074)
9	rs1410051	*TXN*	11/82/353	10/169/462	**0.006**	2.267	(0.925-5.554)	1.440	(0.605-3.428)	**0.635**	**(0.471-0.855)**	**0.038**	**0.763**	**(0.591-0.986)**
10	rs12358630	*ASAH2*	12/162/271	31/187/420	**0.017**	**0.447**	**(0.222-0.899)**	0.600	(0.303-1.189)	**1.343**	**(1.035-1.741)**	0.429	1.090	(0.881-1.348)
11	rs17134961	*C11orf30*	34/209/205	54/245/341	**0.022**	0.738	(0.463-1.178)	1.047	(0.659-1.664)	**1.419**	**(1.102-1.828)**	0.091	1.175	(0.974-1.417)
11	rs2271448	*CAPN1*	15/150/179	34/216/185	**0.012**	0.635	(0.334-1.208)	**0.456**	**(0.24-0.866)**	**0.718**	**(0.536-0.962)**	**0.005**	**0.731**	**(0.586-0.912)**
11	rs16927042	*CD44*	9/100/335	10/102/525	**0.020**	0.918	(0.358-2.354)	1.410	(0.567-3.507)	**1.536**	**(1.129-2.091)**	**0.007**	**1.447**	**(1.106-1.894)**
11	rs3794116	*CD44*	4/96/347	2/105/533	**0.034**	2.188	(0.392-12.214)	3.072	(0.56-16.863)	1.404	(1.032-1.911)	**0.016**	**1.414**	**(1.065-1.878)**
11	rs353632	*CD44*	4/114/326	5/121/508	**0.033**	0.849	(0.222-3.241)	1.247	(0.332-4.677)	1.468	(1.097-1.964)	**0.016**	**1.382**	**(1.063-1.798)**
11	rs16926995	*CD44*	13/90/344	6/128/506	**0.049**	**3.081**	**(1.129-8.412)**	**3.187**	**(1.2-8.466)**	1.034	(0.764-1.399)	0.147	1.214	(0.934-1.578)
11	rs353625	*CD44*	14/122/311	28/212/400	0.050	0.869	(0.441-1.714)	0.643	(0.333-1.242)	0.740	(0.566-0.967)	**0.015**	**0.761**	**(0.61-0.95)**

Chr	SNP	Gene	Case	Control	General	aa vs aA		aa vs AA		aA vs AA		Multiplicative model		
					P value	OR	95% CI	OR	95% CI	OR	95% CI	P value	OR	95% CI
11	rs17571	*CTSD*	0/19/427	1/48/589	**0.034**	NA		NA		0.546	(0.316-0.942)	**0.020**	**0.534**	**(0.312-0.911)**
12	rs607860	*KRT5*	48/192/209	86/313/242	**0.013**	0.910	(0.612-1.352)	**0.646**	**(0.434-0.963)**	**0.710**	**(0.549-0.919)**	**0.006**	**0.776**	**(0.648-0.929)**
12	rs639790	*KRT5*	45/187/215	82/292/266	0.078	0.857	(0.57-1.288)	0.679	(0.453-1.019)	0.792	(0.613-1.024)	**0.024**	**0.811**	**(0.676-0.973)**
12	rs687751	*KRT5*	94/203/151	152/321/166	**0.022**	0.978	(0.716-1.335)	**0.680**	**(0.485-0.954)**	**0.695**	**(0.524-0.922)**	**0.015**	**0.809**	**(0.681-0.96)**
12	rs298120	*KRT6A*	63/196/185	66/271/302	0.071	1.320	(0.893-1.951)	1.558	(1.054-2.303)	1.181	(0.911-1.53)	**0.022**	**1.235**	**(1.031-1.48)**
12	rs406857	*KRT6B*	64/198/184	68/275/296	0.100	1.307	(0.888-1.925)	1.514	(1.027-2.231)	1.158	(0.894-1.501)	**0.034**	**1.215**	**(1.015-1.454)**
14	rs1951133	*CMA1*	46/242/156	79/297/264	**0.032**	0.715	(0.478-1.067)	0.985	(0.651-1.491)	**1.379**	**(1.062-1.79)**	0.326	1.093	(0.915-1.306)
14	rs2073335	*CMA1*	7/119/322	9/129/499	**0.048**	0.843	(0.304-2.335)	1.205	(0.444-3.269)	**1.430**	**(1.074-1.903)**	**0.024**	**1.336**	**(1.039-1.719)**
14	rs1243469	*RNASE7*	12/122/310	13/143/482	0.110	1.082	(0.476-2.459)	1.435	(0.647-3.186)	1.327	(1.002-1.756)	**0.038**	**1.289**	**(1.014-1.639)**
14	rs17094917	*SERPINA12*	53/217/179	104/306/228	0.086	0.719	(0.495-1.044)	0.649	(0.442-0.953)	0.903	(0.695-1.174)	**0.042**	**0.833**	**(0.698-0.993)**
15	rs3743269	*ANXA2*	59/189/199	63/323/250	**0.016**	**1.600**	**(1.075-2.383)**	1.177	(0.788-1.757)	**0.735**	**(0.568-0.952)**	0.645	0.959	(0.801-1.147)
15	rs10444804	*ANXA2*	10/68/370	3/95/543	**0.033**	4.657	(1.235-17.56)	4.892	(1.337-17.896)	1.050	(0.749-1.473)	0.113	1.274	(0.944-1.718)
15	rs894785	*CTSH*	116/233/94	150/308/179	**0.037**	1.022	(0.76-1.375)	**1.473**	**(1.04-2.085)**	**1.441**	**(1.065-1.949)**	**0.030**	**1.210**	**(1.019-1.436)**
15	rs8034542	*CTSH*	36/143/267	55/255/328	**0.018**	1.167	(0.731-1.863)	0.804	(0.513-1.261)	**0.689**	**(0.531-0.894)**	**0.020**	**0.793**	**(0.652-0.964)**
15	rs10400902	*CTSH*	103/231/110	125/317/192	0.104	1.131	(0.828-1.543)	1.438	(1.013-2.042)	1.272	(0.952-1.699)	**0.039**	**1.198**	**(1.009-1.423)**
16	rs2451	*PPL*	31/210/201	84/270/284	**0.004**	**0.474**	**(0.303-0.744)**	**0.521**	**(0.333-0.818)**	1.099	(0.851-1.419)	0.084	0.850	(0.708-1.022)
16	rs2908662	*PPL*	50/234/162	112/293/228	**0.009**	**0.559**	**(0.384-0.813)**	**0.628**	**(0.426-0.928)**	1.124	(0.863-1.465)	0.112	0.867	(0.727-1.034)
17	rs9897526	*GRN*	0/77/367	6/89/544	**0.033**	NA		NA		1.282	(0.92-1.788)	0.522	1.106	(0.812-1.508)
18	rs3910498	*DSC2*	45/186/214	79/296/262	0.068	0.906	(0.602-1.365)	0.697	(0.464-1.049)	0.769	(0.595-0.995)	**0.025**	**0.812**	**(0.676-0.974)**
18	rs276925	*DSC3*	37/200/208	72/303/265	0.117	0.779	(0.504-1.203)	0.655	(0.423-1.013)	0.841	(0.652-1.085)	**0.044**	**0.829**	**(0.69-0.995)**
18	rs276919	*DSC3*	0/44/401	4/85/548	0.051	NA		NA		0.707	(0.481-1.041)	**0.027**	**0.660**	**(0.457-0.956)**
18	rs2290129	*DSG2*	63/259/120	124/322/192	**0.018**	**0.632**	**(0.448-0.891)**	0.813	(0.556-1.188)	1.287	(0.972-1.704)	0.607	0.956	(0.804-1.136)
18	rs8099530	*DSG3*	14/125/304	16/148/472	0.132	1.036	(0.487-2.206)	1.359	(0.654-2.824)	1.311	(0.993-1.732)	**0.049**	**1.266**	**(1.001-1.602)**
19	rs10426	*KLK10*	7/123/313	12/119/506	**0.002**	0.564	(0.215-1.482)	0.943	(0.367-2.421)	**1.671**	**(1.252-2.23)**	**0.004**	**1.447**	**(1.125-1.861)**
19	rs3810091	*KLK13*	12/88/342	12/167/455	**0.040**	1.898	(0.819-4.399)	1.330	(0.59-2.998)	**0.701**	**(0.523-0.94)**	0.116	0.818	(0.637-1.051)
19	rs11665682	*KLK13*	3/13/172	0/9/247	**0.021**	NA		NA		2.074	(0.867-4.961)	**0.005**	**2.974**	**(1.33-6.651)**
19	rs17658926	*KLK14*	6/59/382	1/107/532	**0.016**	**10.881**	**(1.279-92.556)**	**8.356**	**(1.002-69.693)**	0.768	(0.544-1.083)	0.633	0.927	(0.678-1.266)
19	rs10500304	*KLK14*	8/64/373	4/121/516	**0.032**	3.781	(1.097-13.039)	2.767	(0.827-9.256)	0.732	(0.526-1.019)	0.404	0.883	(0.659-1.183)
19	rs10409028	*KLK5*	3/83/359	13/141/486	0.065	0.392	(0.109-1.416)	0.312	(0.088-1.104)	0.797	(0.588-1.08)	**0.030**	**0.741**	**(0.564-0.973)**
20	rs7422035	*CTSA*	4/86/358	2/92/546	**0.037**	2.140	(0.382-11.98)	3.050	(0.556-16.741)	1.426	(1.032-1.969)	**0.015**	**1.446**	**(1.073-1.948)**
20	rs3865535	*TGM3*	0/17/156	0/11/259	**0.026**	NA		NA		2.566	(1.171-5.62)	**0.017**	**2.485**	**(1.15-5.371)**

Analyses done at the general, multiplicative models. Gene name, odds ratio(OR) and 95%confidence interval (CI)of significant SNPs are highlighted in bold. 'a' represents the minor allele and 'A' the major allele

3.3.2 Validation study

3.3.2.1 Population

The 46 genes found associated in the screening study were subjected to a another round of genotyping for validation. At the time of the validation study 7256 samples were ascertained from NUS and KKH cohort. Of them, 6474 were ethnic Chinese (89.22%) similar to the screening population, suggesting a relatively homogenous collection. Based on the quality checks mentioned earlier for the DNA yield and quality, 6363 samples yielded working quality DNA (1.71% loss). Venn diagram showing the overlap of the three allergic disease is shown in Fig 3.2

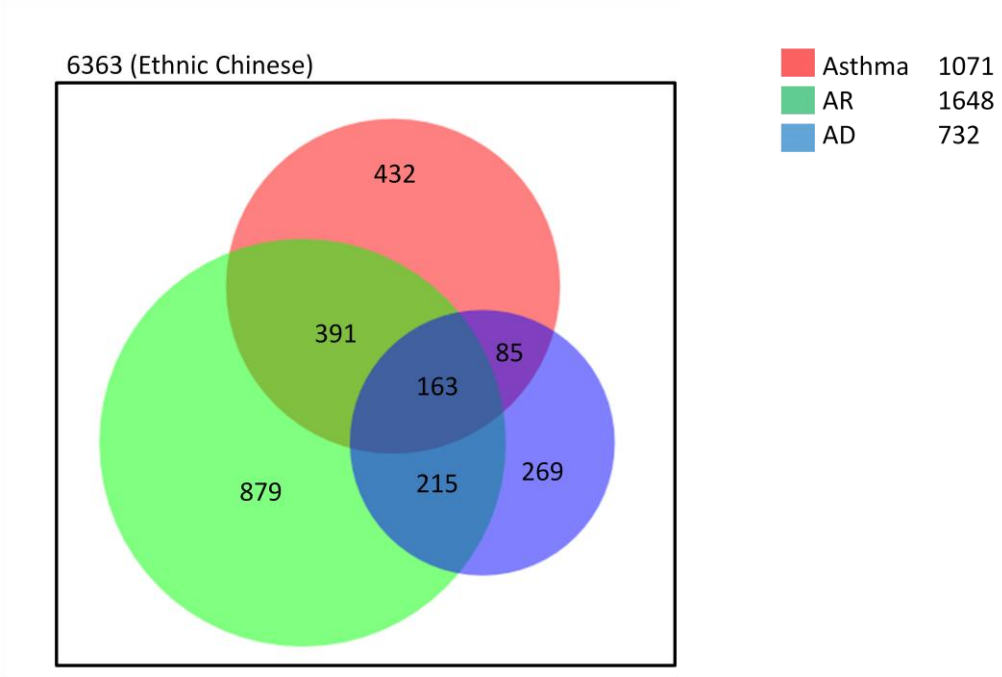

Figure 3-2: Population ascertained for the validation study.
Venn diagram showing the overlapping populations of asthma, allergic rhinitis (AR) and AD. The AD population (732) was genotyped for the validation study in addition to the 413 samples from National Skin Centre (NSC).

By ascertaining 6363 good quality DNA from ethnic Chinese, we collected 1071 (16.8%) asthma cases, 1648 (25.9%) allergic rhinitis cases and 732 (11.5%) AD cases. Asthma proportion were similar to the screening population, but AR and AD saw a slight increase

Stratification of the NUS and KKH cohort with respect to AD is as presented below in Table 3.7 (definitions are as explained in section 3.2.2).

	Atopic Dermatitis	
	#	%
Cases	732	11.50
Controls	1085	17.05
SPT positive	2773	43.58
Phenotype unknown	793	12.46
SPT Inconclusive	685	10.77
Indeterminate	150	2.36
Saline Reaction	145	2.28
Total	6363	100.00

Table 3-7: Ethnic Chinese population's stratification for validation based on the AD questionnaire.

The population stratification is very similar to the screening study indicating continuous homogenous sample collection. The prevalence of cases and controls showed a slight increase as compared to the screening populations with the other categories showing less of a difference.

The validation study was performed on ethnic Chinese samples. In addition to 732 cases and 1085 controls, 650 skin prick positives (AD survey negative)

were also genotyped. The split up based on demographics of the NUS- KKH cohort is shown in Table 3.8.

		Case		Control		SPT positive	
		Number	(%)	Number	(%)	Number	(%)
Gender	Male	336	45.90	327	30.14	374	57.54
	Female	396	54.10	758	69.86	276	42.46
Age group	< 18	62	8.47	11	1.01	76	11.69
	18 - 20	232	31.69	417	38.43	187	28.77
	20.01 - 22	237	32.38	344	31.71	254	39.08
	22.01 - 32	192	26.23	281	25.90	125	19.23
	> 32	9	1.23	32	2.95	8	1.23
Total		732	100	1085	100	650	100

Table 3-8: Demographics of the validation population.
AD cases and controls and skin prick test (SPT) positive stratified according to gender and age group for the validation study.

The sex ratio was again skewed towards the females with the exception of SPT positives. This is due to selection of the 650 highest concentration of SPT samples from a total of 2773 samples. The stratification with respect to age is similar to the screening study. As there is a wide difference within gender and age groups, analyses for the significant SNPs by controlling for gender and age group was performed.

3.3.2.2 Genes and tag SNPs

Tags from the screening study provided the candidate genes for validation. To scrutinize the associations in more detail, the significant genes were re-tagged again. Tagging was performed at r^2 (LD) of 0.8 for SNPs with MAF 0.05 and more. According to CaTS power calculator (Skol *et al.*, 2006) with 732 cases and 1735 controls (1085 +650), genotyping SNPs with MAF of 0.05 can be

found significant at 82% power (1-β error). All genes were re-tagged except *C11orf30* and only the significant tag SNP was re-genotyped on the larger population. This was done due to the SNP number limitation during chip design. The genes tagged for the validation study have been listed below in table 3.9.

Chr	Gene	Sig t- SNPs (Screening)	Genotyped for validation
1	*CA6*	11	21
1	*CAP1*	2	13
1	*CHI3L1*	3	12
1	*FLG2*	1	2
1	*SPRR2G*	2	2
2	*CALM2*	3	5
2	*ZAP70*	1	11
3	*CD80*	2	23
3	*CD86*	1	9
3	*CLDN1*	3	16
5	*GM2A*	1	11
5	*IL7R*	3	7
6	*CDSN*	1	13
6	*DSP*	3	17
6	*OPRM1*	3	41
6	*SERPINB6*	8	11
6	*SMPDL3A*	1	6
8	*ASAH1*	2	20
8	*PLEC1*	3	7
9	*LCN1*	1	2
9	*TXN*	3	14
10	*ASAH2*	1	6
11	*C11orf30*	1	1
11	*CAPN1*	1	7
11	*CD44*	5	45
11	*CTSD*	1	3
12	*KRT5*	3	5
12	*KRT6A*	1	6
12	*KRT6B*	1	4
14	*CMA1*	2	7
14	*RNASE7*	1	5
14	*SERPINA12*	1	20
15	*ANXA2*	2	15
15	*CTSH*	3	7
16	*PPL*	2	13

17	*GRN*	1	8
18	*DSC2*	1	8
18	*DSC3*	2	27
18	*DSG2*	1	10
18	*DSG3*	1	12
19	*KLK10*	1	9
19	*KLK13*	2	6
19	*KLK14*	2	6
19	*KLK5*	1	11
20	*CTSA*	1	5
20	*TGM3*	1	20
Total	46	97	529

Table 3-9: Genes genotyped per chromosome for validation
Chromosome (Chr) wise stratification of the 46 genes represented by 529 tag SNPs that were genotyped in the validation study.

Forty-six genes which were associated at the screening study were re-tagged to give 529 SNPs. These SNPs were genotyped on the larger validation population.

3.3.3 Genotyping results

Analyses for all models of genetic association were tested on the 529 SNPs. SNPs associated at a *P* value of less than 0.05 at either the general or multiplicative models were considered significant. After the analysis, 27 genes remained associated with AD with the maximum number of tags associated from the same gene being *CA6* from Chromosome1. The gene-wide tag SNPs associated are summarised in Table 3.10. Table 3.11 shows the results on the 63 SNPs from 27 genes that remained significant when analysed on 732 cases and 1085 controls.

Chr	Genes	Sig tags
1	*CA6*	8
1	*CAP1*	2
1	*FLG2*	1
1	*SPRR2G*	1
1	*CHI3L1*	4
2	*ZAP70*	1
3	*CD80*	2
3	*CLDN1*	3
5	*GM2A*	1
6	*OPRM1*	3
8	*ASAH1*	2
9	*TXN*	3
9	*LCN1*	1
11	*CD44*	6
11	*CAPN1*	3
11	*C11orf30*	1
12	*KRT6B*	1
12	*KRT5*	2
14	*RNASE7*	2
14	*SERPINA12*	1
15	*CTSH*	3
16	*PPL*	3
18	*DSC3*	1
18	*DSG2*	2
19	*KLK10*	2
19	*KLK14*	3
20	*CTSA*	1
Total	27	63

Table 3-10: Summary of the significant genes and SNPs after validation.
Summary of the 27 genes and the number of tag SNPs found significant (sig) after validation.

Chr	SNP	Gene	General	aa vs aA		aa vs AA		aA vs AA		Multiplicative model			Additive	Dominant/ Recessive model		
			P value	OR	95% CI	OR	95% CI	OR	95% CI	P value	OR	95% CI	P value	P value	OR	95% CI
1	rs1409147	*CA6*	0.107	0.746	(0.37-1.505)	0.620	(0.311-1.236)	0.832	(0.676-1.024)	**0.043**	**0.829**	**(0.693-0.992)**	0.035	NS		
1	rs942969	*CA6*	**0.045**	0.519	(0.249-1.08)	**0.452**	**(0.22-0.93)**	0.872	(0.708-1.074)	**0.038**	**0.825**	**(0.689-0.987)**	0.031	0.037	0.471	(0.229-0.967) Rec
1	rs2274327	*CA6*	**0.009**	0.879	(0.593-1.303)	1.196	(0.812-1.763)	**1.361**	**(1.118-1.658)**	**0.013**	**1.212**	**(1.043-1.409)**	0.012	0.003	1.336	(1.106-1.615) Dom
1	rs2274328	*CA6*	**0.022**	0.938	(0.723-1.218)	1.261	(0.952-1.67)	**1.344**	**(1.086-1.664)**	**0.043**	**1.150**	**(1.005-1.315)**	0.039	0.007	1.322	(1.08-1.619) Dom
1	rs11576766	*CA6*	**0.015**	**0.664**	**(0.469-0.939)**	0.840	(0.594-1.187)	**1.265**	**(1.038-1.541)**	0.609	1.040	(0.902-1.2)	0.585	NS		
1	rs17032921	*CA6*	**0.035**	1.814	(0.958-3.434)	**2.090**	**(1.123-3.889)**	1.152	(0.929-1.428)	**0.026**	**1.236**	**(1.029-1.484)**	0.023	0.026	2.012	(1.084-3.735) Rec
1	rs7545200	*CA6*	0.083	1.041	(0.699-1.552)	1.285	(0.872-1.892)	1.234	(1.01-1.508)	**0.032**	**1.187**	**(1.017-1.386)**	0.033	0.028	1.242	(1.026-1.502) Dom
1	rs3765965	*CA6*	**0.025**	1.788	(0.854-3.744)	**2.190**	**(1.064-4.507)**	1.225	(0.981-1.53)	**0.012**	**1.285**	**(1.059-1.56)**	0.011	0.030	1.276	(1.028-1.583) Dom
1	rs3122413	*CAP1*	**0.045**	1.149	(0.894-1.477)	0.863	(0.655-1.136)	**0.751**	**(0.598-0.942)**	0.206	0.914	(0.796-1.049)	0.200	0.028	0.784	(0.633-0.97) Dom
1	rs16826852	*CAP1*	**0.014**	**1.817**	**(1.209-2.73)**	**1.561**	**(1.053-2.314)**	0.859	(0.702-1.052)	0.604	1.042	(0.891-1.219)	0.608	0.012	1.651	(1.122-2.43) Rec
1	rs2282302	*FLG2*	**0.010**	**2.632**	**(1.382-5.012)**	**2.537**	**(1.352-4.759)**	0.964	(0.777-1.195)	0.227	1.122	(0.936-1.345)	0.206	0.004	2.564	(1.371-4.796) Rec
1	rs513555	*SPRR2G*	**0.002**	1.181	(0.878-1.587)	**1.581**	**(1.17-2.137)**	**1.339**	**(1.093-1.64)**	**0.001**	**1.278**	**(1.113-1.468)**	0.001	0.001	1.388	(1.145-1.682) Dom
1	rs7542294	*CHI3L1*	**0.015**	1.019	(0.772-1.343)	**1.373**	**(1.024-1.841)**	**1.348**	**(1.084-1.676)**	**0.012**	**1.198**	**(1.042-1.378)**	0.011	0.004	1.354	(1.102-1.665) Dom
1	rs7515776	*CHI3L1*	**0.039**	0.864	(0.607-1.231)	1.117	(0.788-1.583)	**1.293**	**(1.061-1.575)**	0.067	1.149	(0.993-1.33)	0.063	0.017	1.261	(1.044-1.522) Dom
1	rs6691378	*CHI3L1*	0.101	1.025	(0.723-1.454)	1.261	(0.893-1.782)	1.230	(1.002-1.509)	**0.048**	**1.163**	**(1.003-1.35)**	0.046	0.033	1.235	(1.017-1.501) Dom
1	rs946261	*CHI3L1*	**0.015**	**0.650**	**(0.471-0.898)**	**0.632**	**(0.458-0.873)**	0.972	(0.796-1.187)	**0.024**	**0.849**	**(0.737-0.978)**	0.025	0.005	0.641	(0.471-0.872) Rec
2	rs1020396	*ZAP70*	**0.033**	0.685	(0.465-1.008)	0.858	(0.586-1.256)	**1.254**	**(1.03-1.526)**	0.403	1.066	(0.919-1.237)	0.396	NS		
3	rs9877854	*CD80*	**0.022**	**2.482**	**(1.288-4.783)**	**2.328**	**(1.231-4.401)**	0.938	(0.752-1.17)	0.360	1.095	(0.907-1.322)	0.347	0.009	2.365	(1.254-4.461) Rec
3	rs6778945	*CD80*	**0.013**	**2.683**	**(1.369-5.256)**	**2.468**	**(1.289-4.726)**	0.920	(0.728-1.163)	0.332	1.107	(0.907-1.35)	0.327	0.005	2.515	(1.316-4.806) Rec
3	rs1060679	*CLDN1*	0.093	0.862	(0.409-1.82)	0.692	(0.333-1.438)	0.802	(0.647-0.994)	**0.034**	**0.815**	**(0.676-0.984)**	0.030	0.034	0.795	(0.644-0.98) Dom
3	rs3774032	*CLDN1*	0.102	0.775	(0.147-4.084)	0.573	(0.111-2.964)	0.739	(0.555-0.985)	**0.033**	**0.745**	**(0.568-0.976)**	0.031	0.035	0.734	(0.553-0.974) Dom
3	rs6776378	*CLDN1*	**0.006**	0.608	(0.328-1.126)	**0.475**	**(0.261-0.864)**	**0.781**	**(0.63-0.969)**	**0.001**	**0.744**	**(0.62-0.893)**	0.002	0.006	0.744	(0.604-0.916) Dom
5	rs153449	*GM2A*	**0.020**	0.792	(0.585-1.073)	1.059	(0.778-1.441)	**1.337**	**(1.084-1.648)**	0.177	1.106	(0.958-1.275)	0.173	0.021	1.268	(1.04-1.546) Dom
6	rs9383692	*OPRM1*	**0.043**	0.793	(0.628-1.003)	**0.721**	**(0.553-0.94)**	0.909	(0.724-1.14)	**0.016**	**0.848**	**(0.742-0.969)**	0.016	0.020	0.767	(0.615-0.957) Rec
6	rs9479769	*OPRM1*	**0.043**	**0.659**	**(0.468-0.929)**	**0.664**	**(0.472-0.933)**	1.007	(0.826-1.227)	0.094	0.884	(0.766-1.02)	0.092	0.014	0.662	(0.477-0.917) Rec
6	rs790258	*OPRM1*	**0.048**	1.982	(0.868-4.524)	**2.339**	**(1.043-5.244)**	1.180	(0.939-1.484)	**0.032**	**1.252**	**(1.023-1.533)**	0.029	NS		
8	rs420610	*ASAH1*	**0.012**	**1.431**	**(1.13-1.812)**	**1.310**	**(1.008-1.702)**	0.916	(0.731-1.147)	0.056	1.141	(0.998-1.305)	0.058	0.004	1.384	(1.11-1.726) Rec
8	rs7460992	*ASAH1*	0.113	1.230	(0.95-1.592)	1.337	(1.017-1.759)	1.087	(0.879-1.345)	**0.049**	**1.147**	**(1.002-1.312)**	0.047	NS		
9	rs4135188	*TXN*	**0.015**	1.155	(0.876-1.523)	0.852	(0.642-1.13)	**0.737**	**(0.6-0.907)**	0.065	0.877	(0.765-1.006)	0.063	0.007	0.765	(0.63-0.928) Dom
9	rs10980288	*TXN*	0.059	1.352	(0.37-4.942)	0.946	(0.266-3.368)	0.700	(0.518-0.946)	**0.033**	**0.736**	**(0.557-0.972)**	0.031	0.024	0.710	(0.529-0.953) Dom
9	rs4135168	*TXN*	**0.043**	**9.704**	**(1.147-82.072)**	**8.805**	**(1.057-73.327)**	0.907	(0.675-1.22)	0.834	1.033	(0.785-1.359)	0.817	0.020	8.912	(1.071-74.193) Rec
9	rs13284515	*LCN1*	**0.016**	0.489	(0.184-1.3)	0.675	(0.258-1.768)	**1.382**	**(1.097-1.742)**	**0.041**	**1.245**	**(1.012-1.533)**	0.037	0.014	1.335	(1.064-1.675) Dom

Table 3-11: Analyses for the significant tag SNPs after validation.

Chr	SNP	Gene	General	aa vs aA		aa vs AA		aA vs AA		Multiplicative model			Additive	Dominant/ Recessive model		
			P value	OR	95% CI	OR	95% CI	OR	95% CI	P value	OR	95% CI	P value	P value	OR	95% CI
11	rs353625	*CD44*	0.085	1.082	(0.649-1.802)	0.858	(0.523-1.408)	0.794	(0.646-0.976)	**0.050**	**0.843**	**(0.712-0.998)**	0.048	0.031	0.801	(0.657-0.976) Dom
11	rs3794116	*CD44*	**0.010**	2.416	(0.808-7.218)	**3.161**	**(1.075-9.296)**	**1.309**	**(1.031-1.66)**	**0.005**	**1.363**	**(1.099-1.692)**	0.004	0.011	1.357	(1.074-1.714) Dom
11	rs353632	*CD44*	**0.030**	1.203	(0.528-2.741)	1.587	(0.708-3.559)	**1.319**	**(1.056-1.648)**	**0.011**	**1.296**	**(1.064-1.578)**	0.009	0.010	1.333	(1.072-1.657) Dom
11	rs16927042	*CD44*	**0.011**	1.233	(0.577-2.633)	1.712	(0.82-3.574)	**1.388**	**(1.095-1.761)**	**0.003**	**1.381**	**(1.121-1.701)**	0.003	0.004	1.411	(1.121-1.776) Dom
11	rs7934770	*CD44*	**0.021**	2.000	(0.939-4.261)	1.499	(0.718-3.132)	**0.749**	**(0.595-0.944)**	0.129	0.852	(0.697-1.042)	0.117	0.038	0.787	(0.629-0.984) Dom
11	rs713330	*CD44*	**0.031**	1.349	(0.371-4.911)	0.931	(0.262-3.313)	**0.690**	**(0.521-0.914)**	**0.017**	**0.727**	**(0.56-0.943)**	0.016	0.012	0.698	(0.529-0.921) Dom
11	rs3825072	*CAPN1*	**0.011**	0.865	(0.581-1.287)	**0.666**	**(0.451-0.982)**	**0.770**	**(0.632-0.939)**	**0.003**	**0.794**	**(0.682-0.925)**	0.003	0.004	0.754	(0.624-0.911) Dom
11	rs17146724	*CAPN1*	0.106	1.201	(0.782-1.843)	1.416	(0.93-2.156)	1.180	(0.968-1.437)	**0.040**	**1.177**	**(1.009-1.373)**	0.034	NS		
11	rs2271448	*CAPN1*	**0.002**	0.727	(0.445-1.189)	**0.527**	**(0.324-0.857)**	**0.725**	**(0.58-0.906)**	**0.001**	**0.744**	**(0.626-0.884)**	0.000	0.001	0.698	(0.562-0.866) Dom
11	rs17134961	*C11orf30*	**0.019**	0.785	(0.549-1.123)	1.039	(0.728-1.483)	**1.324**	**(1.085-1.616)**	0.099	1.134	(0.979-1.312)	0.090	0.013	1.272	(1.051-1.539) Dom
12	rs11170124	*KRT6B*	0.068	1.091	(0.193-6.181)	0.720	(0.132-3.945)	0.660	(0.456-0.957)	**0.028**	**0.676**	**(0.477-0.959)**	0.030	0.027	0.663	(0.461-0.953) Dom
12	rs687751	*KRT5*	**0.029**	1.106	(0.869-1.407)	0.820	(0.63-1.067)	**0.741**	**(0.595-0.924)**	0.095	0.891	(0.779-1.018)	0.093	0.013	0.766	(0.623-0.941) Dom
12	rs607860	*KRT5*	0.060	1.033	(0.769-1.389)	0.815	(0.604-1.098)	0.789	(0.645-0.965)	**0.049**	**0.869**	**(0.757-0.998)**	0.046	0.019	0.794	(0.656-0.961) Dom
14	rs2765901	*RNASE7*	**0.004**	0.757	(0.554-1.034)	**0.609**	**(0.445-0.834)**	**0.805**	**(0.659-0.983)**	**0.001**	**0.790**	**(0.686-0.908)**	0.001	0.005	0.760	(0.628-0.919) Dom
14	rs1243469	*RNASE7*	**0.023**	1.156	(0.609-2.195)	1.529	(0.821-2.85)	**1.322**	**(1.062-1.646)**	**0.007**	**1.300**	**(1.078-1.569)**	0.006	0.008	1.339	(1.084-1.655) Dom
14	rs17094917	*SERPINA12*	0.076	0.827	(0.617-1.109)	0.713	(0.527-0.965)	0.862	(0.698-1.065)	**0.026**	**0.851**	**(0.739-0.98)**	0.024	NS		
15	rs894785	*CTSH*	**0.022**	0.957	(0.761-1.204)	**1.323**	**(1.009-1.735)**	**1.383**	**(1.093-1.749)**	**0.049**	**1.145**	**(1.002-1.309)**	0.045	0.006	1.363	(1.09-1.703) Dom
15	rs8034542	*CTSH*	**0.017**	1.098	(0.755-1.595)	0.823	(0.573-1.184)	**0.750**	**(0.614-0.916)**	**0.016**	**0.830**	**(0.714-0.965)**	0.017	0.005	0.762	(0.631-0.921) Dom
15	rs10400902	*CTSH*	0.060	1.054	(0.832-1.334)	1.339	(1.022-1.755)	1.271	(1.011-1.597)	**0.032**	**1.157**	**(1.013-1.322)**	0.030	0.022	1.292	(1.041-1.603) Dom
16	rs2734741	*PPL*	**0.022**	**1.683**	**(1.16-2.441)**	**1.505**	**(1.045-2.169)**	0.895	(0.734-1.09)	0.428	1.062	(0.916-1.231)	0.422	0.013	1.581	(1.109-2.255) Rec
16	rs2908662	*PPL*	**0.002**	**0.605**	**(0.455-0.805)**	**0.705**	**(0.525-0.948)**	1.165	(0.949-1.432)	0.164	0.906	(0.79-1.038)	0.157	0.001	0.647	(0.493-0.848) Rec
16	rs2451	*PPL*	**0.002**	**0.553**	**(0.393-0.777)**	**0.602**	**(0.428-0.846)**	1.089	(0.893-1.328)	0.089	0.881	(0.764-1.017)	0.083	0.001	0.577	(0.416-0.8) Rec
18	rs8089682	*DSC3*	0.063	0.912	(0.481-1.73)	1.186	(0.637-2.209)	1.301	(1.041-1.626)	**0.035**	**1.233**	**(1.019-1.491)**	0.033	0.023	1.290	(1.04-1.601) Dom
18	rs9304098	*DSG2*	0.057	1.276	(0.955-1.704)	1.428	(1.065-1.914)	1.119	(0.913-1.372)	**0.021**	**1.179**	**(1.027-1.354)**	0.021	0.034	1.346	(1.025-1.769) Rec
18	rs2290129	*DSG2*	**0.008**	**0.695**	**(0.537-0.9)**	0.889	(0.666-1.185)	**1.278**	**(1.025-1.593)**	0.759	0.978	(0.856-1.118)	0.740	0.027	0.756	(0.59-0.968) Rec
19	rs10426	*KLK10*	**0.001**	0.691	(0.316-1.512)	1.057	(0.491-2.273)	**1.529**	**(1.221-1.915)**	**0.001**	**1.384**	**(1.135-1.686)**	0.001	0.000	1.493	(1.199-1.859) Dom
19	rs10425377	*KLK10*	0.078	1.078	(0.746-1.559)	1.319	(0.919-1.892)	1.223	(1.003-1.491)	**0.029**	**1.183**	**(1.019-1.373)**	0.028	0.027	1.238	(1.025-1.495) Dom
19	rs2569488	*KLK14*	**0.037**	1.459	(0.597-3.569)	1.057	(0.443-2.524)	**0.724**	**(0.563-0.932)**	**0.033**	**0.782**	**(0.624-0.98)**	0.034	0.017	0.742	(0.581-0.948) Dom
19	rs6509518	*KLK14*	0.108	0.909	(0.441-1.875)	0.727	(0.359-1.473)	0.799	(0.641-0.997)	**0.037**	**0.815**	**(0.672-0.988)**	0.036	0.039	0.794	(0.64-0.985) Dom
19	rs10500304	*KLK14*	**0.021**	1.802	(0.742-4.373)	1.282	(0.541-3.038)	**0.712**	**(0.553-0.916)**	**0.043**	**0.790**	**(0.63-0.99)**	0.041	0.017	0.739	(0.579-0.945) Dom
20	rs742035	*CTSA*	**0.013**	2.865	(0.547-15.019)	3.897	(0.754-20.151)	**1.360**	**(1.052-1.758)**	**0.008**	**1.389**	**(1.093-1.765)**	0.006	0.012	1.394	(1.081-1.796) Dom

Analyses done at the general, multiplicative, additive, dominant (Dom) and recessive (Rec) model. Gene name, odds ratio(OR) and 95%confidence interval (CI)of significant SNPs are highlighted in bold. 'a' represents the minor allele and 'A' the major allele.

In addition, analysis was also performed on 732 cases and 1735 controls (1085 AD controls + 650 SPT positive), and 732 cases and 650 SPT positives. Results for the significant SNPs shown in table 3.12 are below.

Cases were further subdivided into chronic, non-chronic and mild, moderate and severe cases as mentioned in the section 3.2.2. Table 3.13 summarises the stratification of cases.

Analyses were done between chronic versus controls, non-chronic versus controls, mild/moderate versus controls and severe versus controls. The results did not seem to add more value than the case versus control analyses. To further tease out the effect of SNPs on chronicity and severity, analyses were performed within the different categories of cases. Table 3.14 shows the significant leads of SNPs differentiating chronic cases (n=269) from the non-chronic ones (n=452). However, when a similar approach was adopted for severe cases (n=92) versus mild cases (n=428), the unbalanced analyses did not yield conclusive results.

			Case vs Control + SPT positive (732 vs 1735)								Case vs SPT positive (732 vs 650)							
			General	aa vs aA		aa vs AA		aA vs AA		Multiplicative model	General	aa vs aA		aa vs AA		aA vs AA		Multiplicative model
Chr	SNP	Gene	P value	OR	95% CI	OR	95% CI	OR	95% CI	P value OR 95% CI	P value	OR	95% CI	OR	95% CI	OR	95% CI	P value OR 95% CI
1	rs2274327	*CA6*	**0.023**	0.918	(0.639-1.319)	1.181	(0.826-1.688)	1.287	(1.073-1.542)	**0.022** 1.175 (1.024-1.348)	0.347	0.985	(0.63-1.539)	1.155	(0.744-1.794)	1.173	(0.94-1.463)	0.200 1.117 (0.943-1.321)
1	rs2274328	*CA6*	**0.016**	0.961	(0.756-1.221)	1.274	(0.983-1.652)	1.327	(1.089-1.616)	**0.025** 1.152 (1.018-1.304)	0.080	1.000	(0.743-1.346)	1.298	(0.944-1.785)	1.298	(1.022-1.649)	0.061 1.156 (0.993-1.345)
1	rs11576766	*CA6*	**0.013**	0.704	(0.509-0.974)	0.892	(0.645-1.234)	1.267	(1.057-1.519)	0.372 1.062 (0.931-1.212)	0.085	0.782	(0.525-1.163)	0.994	(0.669-1.476)	1.271	(1.019-1.587)	0.252 1.099 (0.935-1.293)
1	rs3765965	*CA6*	0.124	1.316	(0.712-2.432)	1.548	(0.852-2.812)	1.176	(0.96-1.441)	**0.043** 1.198 (1.005-1.427)	0.755	0.933	(0.46-1.893)	1.027	(0.518-2.036)	1.100	(0.858-1.41)	0.519 1.073 (0.867-1.327)
1	rs3131681	*CAP1*	**0.035**	1.075	(0.858-1.348)	0.825	(0.644-1.056)	0.767	(0.627-0.939)	0.083 0.897 (0.793-1.014)	0.052	1.080	(0.82-1.421)	0.793	(0.584-1.076)	0.734	(0.571-0.944)	0.094 0.879 (0.756-1.022)
1	rs3131687	*CAP1*	**0.021**	1.094	(0.874-1.368)	0.821	(0.643-1.05)	0.751	(0.613-0.92)	0.077 0.895 (0.791-1.012)	**0.034**	1.073	(0.818-1.408)	0.771	(0.57-1.044)	0.719	(0.559-0.925)	0.065 0.868 (0.747-1.009)
1	rs3122413	*CAP1*	**0.009**	1.137	(0.902-1.434)	0.820	(0.636-1.056)	0.721	(0.585-0.889)	0.076 0.891 (0.784-1.012)	**0.011**	1.117	(0.842-1.482)	0.748	(0.545-1.027)	0.670	(0.514-0.872)	**0.047** 0.853 (0.729-0.998)
1	rs16826852	*CAP1*	**0.004**	1.857	(1.282-2.69)	1.668	(1.167-2.384)	0.898	(0.745-1.084)	0.277 1.083 (0.938-1.252)	**0.024**	1.932	(1.186-3.147)	1.871	(1.166-3)	0.968	(0.771-1.217)	0.109 1.158 (0.968-1.385)
1	rs3131676	*CAP1*	0.243	1.216	(0.928-1.595)	1.263	(0.959-1.664)	1.039	(0.859-1.256)	0.147 1.100 (0.967-1.252)	0.118	1.264	(0.896-1.784)	1.434	(1.012-2.031)	1.134	(0.899-1.431)	**0.042** 1.180 (1.006-1.385)
1	rs7542294	*CHI3L1*	**0.024**	1.001	(0.777-1.29)	1.308	(0.998-1.714)	1.307	(1.068-1.599)	**0.018** 1.167 (1.026-1.328)	0.230	0.973	(0.711-1.33)	1.202	(0.862-1.676)	1.236	(0.964-1.585)	0.177 1.116 (0.952-1.308)
1	rs2282302	*FLG2*	**0.016**	2.168	(1.266-3.715)	1.963	(1.163-3.314)	0.905	(0.743-1.104)	0.604 1.045 (0.886-1.232)	0.116	1.705	(0.89-3.265)	1.390	(0.738-2.617)	0.815	(0.641-1.037)	0.482 0.931 (0.762-1.136)
1	rs513555	*SPRR2G*	**0.039**	1.176	(0.9-1.535)	1.392	(1.058-1.831)	1.184	(0.981-1.429)	**0.012** 1.177 (1.037-1.336)	0.642	1.168	(0.844-1.617)	1.119	(0.799-1.567)	0.958	(0.761-1.207)	0.706 1.030 (0.883-1.202)
2	rs1020396	*ZAP70*	**0.036**	0.694	(0.484-0.994)	0.843	(0.591-1.203)	1.215	(1.015-1.456)	0.516 1.046 (0.913-1.2)	0.202	0.709	(0.462-1.09)	0.819	(0.536-1.251)	1.155	(0.925-1.441)	0.868 1.014 (0.858-1.199)
2	rs6543040	*ZAP70*	0.137	0.896	(0.584-1.375)	0.762	(0.502-1.156)	0.851	(0.707-1.024)	**0.048** 0.861 (0.742-0.999)	0.114	0.893	(0.537-1.484)	0.723	(0.441-1.187)	0.810	(0.647-1.015)	**0.040** 0.829 (0.693-0.992)
3	rs9877854	*CD80*	**0.009**	2.381	(1.337-4.241)	2.308	(1.321-4.033)	0.969	(0.79-1.189)	0.195 1.121 (0.943-1.333)	0.084	2.221	(1.033-4.774)	2.276	(1.082-4.787)	1.025	(0.797-1.317)	0.159 1.167 (0.941-1.448)
3	rs6778945	*CD80*	**0.003**	2.732	(1.503-4.968)	2.546	(1.432-4.527)	0.932	(0.75-1.158)	0.219 1.122 (0.934-1.347)	**0.038**	2.825	(1.232-6.481)	2.692	(1.202-6.029)	0.953	(0.729-1.245)	0.238 1.148 (0.913-1.444)
3	rs16829984	*CD80*	0.164	0.850	(0.6-1.204)	0.752	(0.534-1.06)	0.884	(0.738-1.06)	0.061 0.878 (0.766-1.006)	**0.034**	0.899	(0.594-1.36)	0.693	(0.46-1.045)	0.771	(0.618-0.962)	**0.013** 0.812 (0.689-0.957)
3	rs3804588	*CD86*	**0.015**	0.537	(0.111-2.589)	0.601	(0.127-2.839)	1.120	(0.839-1.495)	0.641 1.067 (0.812-1.401)	0.763	0.560	(0.091-3.457)	0.598	(0.1-3.591)	1.068	(0.75-1.52)	0.892 1.023 (0.734-1.426)
3	rs3792287	*CD86*	0.063	0.843	(0.304-2.334)	0.648	(0.238-1.766)	0.769	(0.61-0.969)	**0.021** 0.780 (0.632-0.963)	**0.044**	0.683	(0.217-2.145)	0.498	(0.162-1.532)	0.729	(0.552-0.963)	**0.013** 0.729 (0.567-0.937)
3	rs10513846	*CLDN1*	0.090	1.224	(0.984-1.524)	1.296	(1.016-1.655)	1.059	(0.862-1.3)	**0.043** 1.135 (1.004-1.283)	0.217	1.152	(0.879-1.511)	1.304	(0.968-1.758)	1.132	(0.882-1.452)	0.077 1.145 (0.985-1.331)
3	rs3774032	*CLDN1*	0.958	0.608	(0.132-2.801)	0.410	(0.091-1.855)	0.674	(0.517-0.879)	**0.002** 0.673 (0.524-0.864)	**<0.001**	0.469	(0.092-2.384)	0.274	(0.055-1.364)	0.584	(0.43-0.795)	**<0.001** 0.578 (0.434-0.77)
3	rs6776378	*CLDN1*	**0.005**	0.657	(0.365-1.183)	0.511	(0.288-0.904)	0.777	(0.637-0.948)	**0.001** 0.753 (0.635-0.892)	**0.040**	0.758	(0.379-1.517)	0.584	(0.297-1.146)	0.770	(0.605-0.979)	**0.011** 0.768 (0.626-0.942)
3	rs9839711	*CLDN1*	0.962	0.766	(0.249-2.354)	0.603	(0.199-1.825)	0.787	(0.625-0.991)	**0.030** 0.791 (0.64-0.978)	0.076	0.634	(0.181-2.222)	0.479	(0.14-1.647)	0.756	(0.574-0.995)	**0.024** 0.750 (0.583-0.964)
5	rs153449	*GM2A*	**0.011**	0.792	(0.598-1.049)	1.055	(0.793-1.405)	1.332	(1.098-1.617)	0.144 1.103 (0.967-1.258)	0.057	0.792	(0.561-1.118)	1.049	(0.74-1.487)	1.325	(1.044-1.682)	0.256 1.099 (0.934-1.293)
5	rs9324685	*GM2A*	0.378	0.775	(0.53-1.133)	0.771	(0.531-1.118)	0.995	(0.829-1.194)	0.345 0.934 (0.811-1.076)	0.071	0.707	(0.455-1.098)	0.620	(0.403-0.954)	0.876	(0.701-1.095)	**0.031** 0.829 (0.7-0.983)
5	rs11567761	*IL7R*	0.467	1.447	(0.789-2.652)	1.364	(0.757-2.46)	0.943	(0.772-1.152)	0.976 1.003 (0.844-1.191)	**0.014**	4.401	(1.46-13.261)	3.959	(1.33-11.78)	0.900	(0.707-1.145)	0.661 1.049 (0.848-1.296)
6	rs3130554	*CDSN*	0.818	1.014	(0.789-1.304)	0.954	(0.733-1.241)	0.940	(0.776-1.139)	0.636 0.970 (0.856-1.099)	0.062	0.966	(0.717-1.303)	0.742	(0.539-1.022)	0.768	(0.605-0.976)	**0.038** 0.852 (0.732-0.991)
6	rs2239520	*CDSN*	0.609	1.152	(0.849-1.562)	1.082	(0.796-1.472)	0.940	(0.779-1.133)	0.972 1.002 (0.878-1.144)	**0.040**	1.041	(0.727-1.49)	0.776	(0.537-1.121)	0.746	(0.592-0.94)	**0.039** 0.845 (0.721-0.992)
6	rs3799526	*DSP*	0.252	0.753	(0.522-1.084)	0.739	(0.515-1.06)	0.982	(0.82-1.176)	0.231 0.920 (0.802-1.055)	**0.027**	0.625	(0.412-0.95)	0.571	(0.378-0.864)	0.913	(0.732-1.14)	**0.025** 0.828 (0.702-0.976)

Table 3-12: Analyses for the significant tag SNPs from Case vs Controls+ SPT positives and Case vs SPT positives.

			Case vs Control + SPT positive (732 vs 1735)										Case vs SPT positive (732 vs 650)									
			General	aa vs aA		aa vs AA		aA vs AA		Multiplicative model			General	aa vs aA		aa vs AA		aA vs AA		Multiplicative model		
Chr	SNP	Gene	P value	OR	95% CI	OR	95% CI	OR	95% CI	P value	OR	95% CI	P value	OR	95% CI	OR	95% CI	OR	95% CI	P value	OR	95% CI
6	rs4870266	*OPRM1*	0.264	0.645	(0.355-1.173)	0.717	(0.399-1.287)	1.111	(0.916-1.348)	0.794	1.022	(0.867-1.205)	**0.040**	1.219	(0.52-2.857)	1.627	(0.705-3.752)	1.335	(1.048-1.7)	**0.013**	1.305	(1.056-1.612)
6	rs318426	*SERPINB6*	**0.006**	NA		NA		0.940	(0.629-1.405)	0.767	0.942	(0.633-1.4)	0.340	NA		NA		0.788	(0.493-1.259)	0.325	0.793	(0.499-1.259)
6	rs2295767	*SERPINB6*	0.119	1.107	(0.685-1.79)	1.328	(0.835-2.112)	1.200	(0.986-1.46)	**0.037**	1.187	(1.01-1.396)	0.190	1.143	(0.618-2.112)	1.392	(0.769-2.519)	1.218	(0.953-1.555)	0.065	1.210	(0.988-1.482)
6	rs13192569	*SMPDL3A*	**0.006**	0.453	(0.248-0.829)	0.573	(0.317-1.036)	1.265	(1.043-1.534)	0.407	1.072	(0.91-1.262)	**0.001**	0.361	(0.182-0.717)	0.532	(0.273-1.036)	1.474	(1.154-1.883)	0.162	1.157	(0.943-1.419)
6	rs11751996	*SMPDL3A*	0.251	0.613	(0.308-1.22)	0.579	(0.296-1.132)	0.945	(0.762-1.172)	0.188	0.882	(0.731-1.064)	**0.032**	0.496	(0.231-1.066)	0.417	(0.199-0.874)	0.840	(0.647-1.091)	**0.017**	0.762	(0.61-0.953)
8	rs420610	*ASAH1*	**0.032**	1.330	(1.073-1.649)	1.178	(0.927-1.496)	0.885	(0.718-1.092)	0.202	1.084	(0.958-1.226)	0.263	1.190	(0.918-1.543)	0.993	(0.74-1.332)	0.834	(0.645-1.08)	0.951	0.995	(0.856-1.157)
8	rs6558408	*PLEC1*	0.083	1.362	(1.038-1.787)	1.276	(0.972-1.676)	0.937	(0.778-1.129)	0.288	1.072	(0.943-1.219)	**0.036**	1.527	(1.084-2.149)	1.269	(0.898-1.794)	0.831	(0.663-1.043)	0.714	1.030	(0.88-1.204)
9	rs13284515	*LCN1*	**0.008**	0.462	(0.185-1.152)	0.628	(0.255-1.544)	1.359	(1.101-1.677)	**0.041**	1.218	(1.008-1.471)	0.052	0.425	(0.15-1.199)	0.561	(0.202-1.554)	1.321	(1.017-1.717)	0.179	1.174	(0.929-1.483)
9	rs4135192	*TXN*	**0.021**	0.753	(0.237-2.396)	0.782	(0.251-2.436)	1.039	(0.802-1.345)	0.934	1.010	(0.795-1.283)	0.893	1.222	(0.266-5.614)	1.130	(0.252-5.073)	0.924	(0.676-1.264)	0.694	0.943	(0.706-1.261)
9	rs4135188	*TXN*	**0.010**	1.164	(0.901-1.503)	0.868	(0.669-1.126)	0.746	(0.617-0.902)	0.060	0.886	(0.78-1.005)	0.067	1.179	(0.863-1.611)	0.897	(0.651-1.236)	0.761	(0.603-0.96)	0.183	0.900	(0.772-1.051)
9	rs10980288	*TXN*	0.055	1.244	(0.379-4.089)	0.897	(0.28-2.87)	0.721	(0.543-0.956)	**0.031**	0.751	(0.579-0.975)	0.286	1.083	(0.261-4.493)	0.822	(0.205-3.305)	0.759	(0.54-1.068)	0.119	0.779	(0.568-1.067)
11	rs3825072	*CAPN1*	**0.029**	0.930	(0.639-1.353)	0.743	(0.515-1.072)	0.799	(0.666-0.96)	**0.011**	0.832	(0.722-0.958)	0.359	1.069	(0.671-1.704)	0.909	(0.576-1.435)	0.850	(0.68-1.063)	0.235	0.900	(0.757-1.071)
11	rs17146724	*CAPN1*	0.095	1.074	(0.732-1.577)	1.287	(0.883-1.876)	1.199	(0.999-1.437)	**0.035**	1.164	(1.01-1.34)	0.187	0.908	(0.572-1.442)	1.118	(0.712-1.757)	1.232	(0.985-1.54)	0.135	1.141	(0.96-1.358)
11	rs625750	*CAPN1*	0.154	0.773	(0.373-1.602)	0.505	(0.238-1.071)	0.653	(0.363-1.174)	**0.049**	0.684	(0.469-0.999)	0.637	1.438	(0.49-4.219)	1.026	(0.338-3.116)	0.713	(0.328-1.55)	0.790	0.932	(0.556-1.563)
11	rs16926995	*CD44*	**0.031**	2.501	(1.229-5.088)	2.204	(1.106-4.393)	0.881	(0.708-1.098)	0.929	1.009	(0.833-1.221)	0.134	2.336	(0.932-5.855)	1.988	(0.811-4.872)	0.851	(0.654-1.109)	0.812	0.972	(0.771-1.227)
11	rs3794116	*CD44*	0.052	1.889	(0.785-4.544)	2.257	(0.953-5.345)	1.195	(0.963-1.483)	**0.030**	1.239	(1.02-1.503)	0.712	1.450	(0.513-4.096)	1.504	(0.543-4.169)	1.037	(0.799-1.346)	0.565	1.071	(0.848-1.353)
11	rs16927042	*CD44*	**0.025**	1.211	(0.618-2.373)	1.581	(0.823-3.036)	1.306	(1.053-1.619)	**0.006**	1.303	(1.08-1.573)	0.347	1.179	(0.512-2.715)	1.397	(0.622-3.137)	1.184	(0.91-1.541)	0.139	1.190	(0.945-1.499)
11	rs7934770	*CD44*	**0.048**	1.716	(0.88-3.346)	1.345	(0.704-2.571)	0.784	(0.632-0.972)	0.151	0.872	(0.723-1.051)	0.429	1.354	(0.601-3.05)	1.149	(0.523-2.526)	0.848	(0.652-1.104)	0.404	0.907	(0.722-1.14)
11	rs713330	*CD44*	0.284	1.365	(0.41-4.547)	0.986	(0.302-3.213)	0.722	(0.554-0.941)	**0.026**	0.758	(0.593-0.968)	0.315	1.397	(0.304-6.428)	1.095	(0.244-4.915)	0.784	(0.568-1.081)	0.187	0.818	(0.607-1.103)
11	rs3751031	*CD44*	0.292	0.672	(0.404-1.12)	0.718	(0.436-1.184)	1.069	(0.888-1.286)	0.837	0.984	(0.844-1.147)	**0.028**	0.540	(0.306-0.953)	0.484	(0.278-0.843)	0.896	(0.715-1.122)	**0.029**	0.816	(0.68-0.979)
11	rs7126359	*CD44*	0.291	0.672	(0.398-1.135)	0.729	(0.437-1.215)	1.084	(0.898-1.309)	0.959	0.996	(0.851-1.166)	**0.045**	0.535	(0.298-0.958)	0.495	(0.281-0.874)	0.926	(0.736-1.165)	0.056	0.833	(0.691-1.005)
11	rs11607862	*CD44*	0.710	1.117	(0.854-1.462)	1.105	(0.844-1.447)	0.989	(0.82-1.192)	0.595	1.035	(0.911-1.177)	**0.045**	1.547	(1.089-2.199)	1.487	(1.045-2.117)	0.961	(0.767-1.205)	0.126	1.131	(0.966-1.325)
11	rs13347	*CD44*	0.100	0.730	(0.536-0.994)	0.723	(0.531-0.983)	0.991	(0.82-1.198)	0.102	0.893	(0.78-1.023)	**0.047**	0.648	(0.449-0.933)	0.650	(0.451-0.934)	1.003	(0.792-1.27)	0.074	0.860	(0.728-1.015)
11	rs4755392	*CD44*	0.172	1.216	(0.972-1.522)	1.232	(0.96-1.58)	1.013	(0.819-1.253)	0.112	1.108	(0.976-1.258)	**0.035**	1.454	(1.094-1.934)	1.338	(0.974-1.837)	0.920	(0.709-1.193)	0.102	1.140	(0.974-1.333)
12	rs687751	*KRT5*	**0.049**	1.141	(0.913-1.427)	0.889	(0.698-1.132)	0.779	(0.637-0.953)	0.230	0.927	(0.82-1.049)	0.258	1.208	(0.916-1.594)	1.018	(0.755-1.374)	0.843	(0.659-1.077)	0.925	0.993	(0.854-1.155)
12	rs1054122	*KRT6A*	0.557	NA		NA		NA		**0.033**	NA		1.000	NA		NA		NA		0.185	NA	
12	rs3907935	*KRT6A*	0.097	1.431	(0.91-2.25)	1.187	(0.766-1.84)	0.829	(0.682-1.008)	0.361	0.929	(0.794-1.088)	**0.033**	1.960	(1.059-3.627)	1.547	(0.847-2.823)	0.789	(0.622-1.001)	0.533	0.940	(0.775-1.141)
14	rs1951133	*CMA1*	0.097	0.782	(0.592-1.033)	0.923	(0.693-1.23)	1.180	(0.979-1.423)	0.778	1.018	(0.897-1.156)	**0.040**	0.673	(0.484-0.936)	0.813	(0.579-1.139)	1.207	(0.959-1.518)	0.750	0.975	(0.835-1.138)
14	rs2765901	*RNASE7*	**0.044**	0.802	(0.597-1.076)	0.699	(0.521-0.939)	0.872	(0.726-1.048)	**0.013**	0.849	(0.745-0.966)	0.818	0.898	(0.624-1.291)	0.892	(0.621-1.282)	0.994	(0.794-1.244)	0.631	0.962	(0.819-1.128)

Chr	SNP	Gene	Case vs Control + SPT positive (732 vs 1735)								Case vs SPT positive (732 vs 650)							
			General	aa vs aA		aa vs AA		aA vs AA		Multiplicative model	General	aa vs aA		aa vs AA		aA vs AA		Multiplicative model
			P value	OR	95% CI	OR	95% CI	OR	95% CI	P value OR 95% CI	P value	OR	95% CI	OR	95% CI	OR	95% CI	P value OR 95% CI
14	rs1243469	*RNASE7*	**0.045**	1.053	(0.597-1.858)	1.337	(0.771-2.321)	1.271	(1.041-1.551)	**0.014** 1.237 (1.044-1.467)	0.374	0.924	(0.468-1.825)	1.100	(0.57-2.126)	1.190	(0.932-1.521)	0.206 1.144 (0.929-1.409)
14	rs17094917	*SERPINA12*	0.103	0.824	(0.627-1.082)	0.740	(0.559-0.979)	0.898	(0.739-1.091)	**0.037** 0.870 (0.764-0.992)	0.372	0.817	(0.586-1.138)	0.786	(0.559-1.106)	0.963	(0.757-1.224)	0.222 0.905 (0.77-1.062)
14	rs2402475	*SERPINA12*	**0.019**	1.042	(0.844-1.286)	1.372	(1.072-1.756)	1.316	(1.061-1.633)	**0.012** 1.171 (1.036-1.324)	**0.011**	1.078	(0.829-1.401)	1.515	(1.122-2.045)	1.406	(1.086-1.819)	**0.006** 1.233 (1.062-1.433)
14	rs6575436	*SERPINA12*	0.165	0.897	(0.658-1.223)	0.782	(0.574-1.065)	0.871	(0.726-1.046)	0.063 0.882 (0.773-1.007)	**0.044**	0.884	(0.612-1.277)	0.695	(0.48-1.005)	0.786	(0.628-0.984)	**0.016** 0.822 (0.701-0.964)
14	rs11951017	*SERPINA12*	0.235	0.909	(0.709-1.167)	0.807	(0.621-1.05)	0.888	(0.733-1.076)	0.095 0.899 (0.794-1.019)	0.062	0.936	(0.694-1.263)	0.727	(0.528-1.003)	0.777	(0.612-0.986)	**0.031** 0.846 (0.727-0.985)
15	rs894785	*CTSH*	**0.015**	0.964	(0.781-1.19)	1.323	(1.029-1.7)	1.372	(1.103-1.707)	**0.031** 1.145 (1.013-1.295)	0.064	0.738	(0.567-0.959)	0.756	(0.558-1.024)	1.025	(0.791-1.329)	0.078 0.874 (0.752-1.015)
15	rs8034542	*CTSH*	**0.047**	1.004	(0.711-1.417)	0.805	(0.577-1.124)	0.802	(0.667-0.965)	**0.022** 0.849 (0.738-0.977)	0.375	0.865	(0.574-1.305)	0.778	(0.524-1.158)	0.900	(0.717-1.129)	0.149 0.882 (0.744-1.046)
15	rs10400902	*CTSH*	**0.035**	0.997	(0.804-1.237)	1.301	(1.014-1.668)	1.304	(1.055-1.612)	**0.032** 1.143 (1.011-1.293)	0.060	0.911	(0.7-1.187)	1.242	(0.921-1.676)	1.363	(1.055-1.76)	0.138 1.120 (0.964-1.302)
16	rs2908662	*PPL*	**0.005**	0.644	(0.493-0.842)	0.732	(0.555-0.966)	1.137	(0.941-1.373)	0.176 0.917 (0.808-1.04)	0.128	0.718	(0.52-0.99)	0.784	(0.561-1.094)	1.092	(0.866-1.376)	0.392 0.935 (0.802-1.091)
16	rs2451	*PPL*	**0.005**	0.586	(0.425-0.808)	0.632	(0.458-0.871)	1.078	(0.899-1.294)	0.097 0.894 (0.784-1.02)	0.084	0.649	(0.442-0.952)	0.688	(0.469-1.009)	1.061	(0.848-1.326)	0.288 0.917 (0.781-1.076)
16	rs2734741	*PPL*	**0.009**	1.688	(1.204-2.367)	1.585	(1.138-2.208)	0.939	(0.783-1.127)	0.172 1.100 (0.96-1.26)	**0.039**	1.697	(1.094-2.631)	1.729	(1.124-2.658)	1.019	(0.816-1.273)	0.072 1.168 (0.986-1.383)
17	rs9897526	*GRN*	**0.004**	0.121	(0.016-0.921)	0.143	(0.019-1.077)	1.179	(0.936-1.485)	0.683 1.046 (0.843-1.297)	**0.026**	0.100	(0.012-0.811)	0.113	(0.014-0.905)	1.131	(0.851-1.502)	0.894 0.983 (0.757-1.275)
17	rs5848	*GRN*	0.081	0.985	(0.736-1.318)	0.807	(0.6-1.085)	0.819	(0.68-0.986)	0.051 0.880 (0.773-1.001)	**0.022**	0.972	(0.686-1.376)	0.718	(0.502-1.028)	0.739	(0.588-0.929)	**0.016** 0.825 (0.706-0.965)
18	rs1586959	*DSC3*	**0.042**	0.584	(0.36-0.946)	0.553	(0.346-0.884)	0.948	(0.782-1.15)	**0.044** 0.850 (0.725-0.996)	0.079	0.555	(0.317-0.971)	0.543	(0.316-0.931)	0.978	(0.771-1.24)	0.122 0.860 (0.709-1.042)
18	rs9304098	*DSG2*	**0.022**	1.298	(0.996-1.692)	1.456	(1.114-1.905)	1.122	(0.929-1.354)	**0.008** 1.188 (1.046-1.349)	0.057	1.338	(0.957-1.87)	1.507	(1.075-2.113)	1.127	(0.896-1.417)	**0.021** 1.204 (1.029-1.408)
18	rs10502572	*DSG2*	0.085	1.185	(0.714-1.965)	1.421	(0.869-2.326)	1.200	(0.992-1.451)	**0.027** 1.195 (1.021-1.4)	0.110	1.091	(0.58-2.054)	1.378	(0.745-2.547)	1.263	(0.997-1.598)	**0.039** 1.231 (1.01-1.499)
18	rs2290129	*DSG2*	**0.001**	0.655	(0.516-0.832)	0.837	(0.641-1.093)	1.278	(1.042-1.567)	0.438 0.952 (0.842-1.077)	**0.001**	0.596	(0.449-0.791)	0.761	(0.555-1.045)	1.278	(0.995-1.641)	0.223 0.911 (0.783-1.058)
18	rs2187047	*DSG2*	0.113	0.914	(0.656-1.274)	1.138	(0.812-1.594)	1.245	(1.013-1.529)	0.126 1.119 (0.969-1.292)	**0.042**	0.825	(0.55-1.237)	1.139	(0.757-1.713)	1.380	(1.073-1.776)	0.101 1.161 (0.972-1.387)
18	rs12959976	*DSG3*	**0.034**	NA		NA		1.332	(0.861-2.058)	0.114 1.404 (0.92-2.143)	0.116	NA		NA		1.620	(0.903-2.908)	0.066 1.702 (0.959-3.022)
18	rs36032521	*DSG3*	0.306	0.839	(0.669-1.052)	0.904	(0.703-1.162)	1.077	(0.879-1.32)	0.519 0.960 (0.849-1.086)	**0.039**	0.724	(0.553-0.948)	0.711	(0.525-0.962)	0.982	(0.762-1.265)	**0.034** 0.850 (0.731-0.988)
18	rs2298614	*DSG3*	0.114	0.968	(0.645-1.453)	1.174	(0.79-1.744)	1.213	(1.007-1.46)	0.058 1.152 (0.995-1.334)	0.061	0.905	(0.547-1.496)	1.192	(0.732-1.941)	1.317	(1.046-1.659)	**0.040** 1.211 (1.009-1.454)
19	rs10426	*KLK10*	**0.006**	0.701	(0.343-1.435)	0.972	(0.482-1.96)	1.386	(1.132-1.697)	**0.008** 1.270 (1.063-1.518)	0.344	0.716	(0.307-1.668)	0.852	(0.372-1.948)	1.189	(0.929-1.522)	0.324 1.115 (0.898-1.383)
19	rs2569488	*KLK14*	**0.049**	1.603	(0.695-3.698)	1.203	(0.533-2.714)	0.751	(0.593-0.95)	0.057 0.814 (0.659-1.006)	0.205	1.947	(0.633-5.989)	1.553	(0.517-4.664)	0.798	(0.6-1.06)	0.306 0.874 (0.676-1.131)
19	rs10500304	*KLK14*	**0.021**	2.055	(0.897-4.707)	1.516	(0.677-3.393)	0.738	(0.582-0.934)	0.073 0.825 (0.668-1.018)	0.092	2.752	(0.839-9.031)	2.158	(0.673-6.922)	0.784	(0.59-1.042)	0.373 0.890 (0.688-1.15)
20	rs742035	*CTSA*	0.078	1.379	(0.429-4.43)	1.762	(0.557-5.573)	1.277	(1.013-1.611)	**0.026** 1.275 (1.029-1.58)	0.592	0.785	(0.221-2.784)	0.908	(0.261-3.153)	1.157	(0.871-1.537)	0.396 1.120 (0.863-1.453)
20	rs3817731	*CTSA*	0.179	NA		NA		1.596	(1.164-2.189)	**0.009** 1.468 (1.099-1.96)	**0.001**	NA		NA		2.048	(1.337-3.137)	**0.002** 1.859 (1.248-2.769)
20	rs6048208	*TGM3*	**0.033**	0.982	(0.248-3.886)	1.019	(0.263-3.952)	1.037	(0.795-1.352)	0.796 1.033 (0.805-1.326)	0.815	1.230	(0.2-7.562)	1.349	(0.225-8.104)	1.097	(0.789-1.525)	0.534 1.103 (0.809-1.505)

Analyses done at the general and multiplicative models. Gene name, odds ratio(OR) and 95% confidence interval (CI)of significant SNPs are highlighted in bold. 'a' represents the minor allele and 'A' the major allele.

		Case	
		Number	(%)
Chronicity	Chronic	269	36.75
	Non-Chronic	452	61.75
	Excluded	11	1.50
Severity	Mild	428	58.47
	Moderate	212	28.96
	Severe	92	12.57
	Total	732	100

Table 3-13: Stratification of cases according to chronicity and severity.

269 chronic vs 452 non-chronic			General	aa vs aA		aa vs AA		aA vs AA		Multiplicative model		
Chr	SNP	Gene	P value	OR	95% CI	OR	95% CI	OR	95% CI	P value	OR	95% CI
1	rs3765964	*CA6*	**0.039**	0.471	(0.258-0.86)	0.482	(0.264-0.879)	1.023	(0.745-1.403)	0.106	0.826	(0.655-1.042)
1	rs11576766	*CA6*	**0.045**	0.506	(0.27-0.947)	0.666	(0.354-1.254)	1.317	(0.961-1.806)	0.868	1.020	(0.811-1.283)
2	rs1020396	*ZAP70*	**0.009**	0.928	(0.474-1.815)	0.582	(0.3-1.129)	0.628	(0.458-0.862)	**0.005**	0.707	(0.554-0.902)
2	rs4851462	*ZAP70*	0.095	0.361	(0.078-1.675)	0.280	(0.061-1.277)	0.776	(0.55-1.094)	**0.046**	0.735	(0.542-0.995)
3	rs1599796	*CD80*	**0.042**	1.980	(1.03-3.806)	2.237	(1.177-4.25)	1.130	(0.823-1.552)	**0.043**	1.285	(1.008-1.637)
3	rs693640	*CD80*	0.092	1.497	(0.865-2.59)	1.796	(1.041-3.098)	1.200	(0.872-1.651)	**0.039**	1.274	(1.012-1.606)
3	rs527004	*CD80*	0.065	1.377	(0.384-4.94)	0.868	(0.251-2.999)	0.630	(0.426-0.932)	**0.037**	0.693	(0.491-0.979)
6	rs11155954	*OPRM1*	**0.020**	0.630	(0.26-1.527)	0.434	(0.182-1.034)	0.689	(0.499-0.951)	**0.006**	0.692	(0.53-0.902)
6	rs9371331	*OPRM1*	**0.047**	0.440	(0.174-1.109)	0.359	(0.144-0.893)	0.816	(0.593-1.121)	**0.030**	0.750	(0.578-0.972)
6	rs7755659	*OPRM1*	0.091	0.840	(0.478-1.478)	0.622	(0.356-1.086)	0.740	(0.533-1.026)	**0.031**	0.769	(0.606-0.977)
6	rs9397685	*OPRM1*	0.117	1.391	(0.687-2.814)	1.823	(0.911-3.647)	1.311	(0.925-1.857)	**0.042**	1.318	(1.01-1.719)
8	rs11986226	*ASAH1*	**0.030**	2.330	(1.225-4.429)	2.129	(1.145-3.96)	0.914	(0.66-1.266)	0.201	1.176	(0.917-1.508)
11	rs625750	*CAPN1*	**0.005**	0.267	(0.071-0.999)	1.160	(0.296-4.539)	4.350	(1.67-11.332)	0.192	1.495	(0.816-2.738)
11	rs7126359	*CD44*	0.121	1.116	(0.445-2.8)	1.537	(0.624-3.784)	1.376	(0.995-1.904)	**0.048**	1.313	(1.002-1.72)
14	rs17094919	*SERPINA12*	**0.027**	2.027	(1.189-3.456)	1.574	(0.932-2.656)	0.776	(0.554-1.086)	0.589	1.068	(0.841-1.355)
14	rs17094917	*SERPINA12*	**0.046**	1.561	(0.966-2.523)	1.069	(0.657-1.74)	0.685	(0.488-0.961)	0.523	0.928	(0.738-1.167)
14	rs2236244	*SERPINA12*	0.091	0.791	(0.271-2.315)	0.547	(0.194-1.54)	0.691	(0.476-1.004)	**0.026**	0.694	(0.503-0.958)
14	rs1951015	*SERPINA12*	0.121	0.711	(0.245-2.058)	0.514	(0.184-1.436)	0.723	(0.498-1.05)	**0.035**	0.710	(0.516-0.977)
18	rs3737375	*DSG2*	**0.006**	1.551	(0.824-2.92)	0.913	(0.49-1.701)	0.589	(0.423-0.819)	**0.039**	0.771	(0.602-0.987)

Table 3-14: Analyses for the significant tag SNPs as markers for chronicity.
Analyses done at the general and multiplicative models. 'a' represents the minor allele and 'A' the major allele. Odds ratio(OR), 95% confidence interval (95% CI).

The wide discrepancy observed in the demographics may form a confounding factor in the case-control association analyses. To verify the associations, the significant SNPs were re-analysed by controlling for gender and age group. The variables used as confounding variables for the genetic analyses are as shown in Table 3.16.

Variable	Odds ratio	(95% CI)	P value
Gender			
Female	Ref	NA	NA
Male	1.874	(1.516-2.316)	<0.001
Age group			
18-20	Ref	NA	NA
<18	7.892	(4.042-15.408)	<0.001
20.01-22	0.994	(0.778-1.270)	0.963
22.01-32	0.989	(0.765-1.279)	0.933
>32	0.431	(0.20-0.926)	0.031

Table 3-15: Variables for used for adjusting genetic association results.

Due to the manner of subject sampling, age was expected to have a significant association with AD as the samples contributing to the age group of less-than-18 years came from the paediatric allergic clinic of KK hospital. Also, gender with higher prevalence of the disease switched from females (chapter 2 for eczema) to males in this cohort of ethnic Chinese AD case-controls. This swap is due to the increased number of males in the less-than-18 years group and above 32 years of age. Sample ascertainment in the paediatric ward resulted in a skew towards males with the odds ratio for males in the group of less than18 years and above 32 years reached 3.38 (1.52- 7.49), where as the age group of

18-32 showed association of males with odds ratio 1.84 (1.50-2.25). To control for any confounding effect that may arise due to age and gender, associated SNPs were re-tested to yield adjusted odds ratio and 95% confidence interval (Table 3.16).

Chr	SNP	Gene	General	Unadjusted		Adjusted for gender and age grp		Multiplicative	Unadjusted		Adj for gender and age grp	
			P value	OR (95% CI)[a,b,c]	P value	OR (95% CI)[a,b,c]		P value	OR (95% CI)	P value	OR (95% CI)	
1	rs1409147	*CA6*	NS		NS			0.043	0.829 (0.693-0.992)	NS		
1	**rs942969**	***CA6***	0.045	0.452 (0.22-0.93)[b]	0.033	0.457 (0.216-0.968)[a], 0.402 (0.192-0.840)[b]		0.038	0.824 (0.689-0.987)	0.02	0.418 (0.20-0.871)	
1	**rs2274327**	***CA6***	0.009	1.361 (1.118-1.658)[c]	0.005	1.405 (1.147-1.723)[c]		0.013	1.212 (1.043-1.409)	0.001	1.376 (1.132-1.673)	
1	**rs2274328**	***CA6***	0.022	1.344 (1.086-1.664)[c]	0.029	0.742 (0.595-0.924)[a]		0.043	1.149 (1.005-1.315)	0.012	0.766 (0.622-0.944)	
1	**rs11576766**	***CA6***	0.015	0.663 (0.469-0.939)[a], 1.264 (1.038-1.541)[c]	0.03	1.556 (1.085-2.231)[c]		NS		0.048	1.418 (1.004-2.003)	
1	**rs17032921**	***CA6***	0.035	2.089 (1.123-3.889)[b]	0.012	1.963 (1.022-3.770)[a], 2.353 (1.246-4.442)[b]		0.026	1.235 (1.029-1.484)	0.03	1.269 (1.024-1.574)	
1	**rs7545200**	***CA6***	NS		NS			0.032	1.187 (1.017-1.386)	0.021	1.262 (1.036-1.537)	
1	**rs3765965**	***CA6***	0.025	2.190 (1.064-4.507)[b]	0.007	0.771 (0.612-0.971)[a], 0.405 (0.195-0.841)[b]		0.012	1.285 (1.059-1.56)	0.024	0.431 (0.208-0.893)	
1	**rs3122413**	***CAP1***	0.045	0.751 (0.598-0.942)[c]	0.039	1.354 (1.071-1.712)[a]		NS		0.023	1.291 (1.035-1.609)	
1	**rs16826852**	***CAP1***	0.014	1.816 (1.209-2.73)[a], 1.560 (1.053-2.314)[b]	0.022	0.627 (0.418-0.940)[b], 0.556 (0.366-0.846)[c]		NS		0.012	0.6 (0.403-0.892)	
1	**rs2282302**	***FLG2***	0.010	2.631 (1.382-5.012)[a], 2.536 (1.352-4.759)[b]	0.047	2.281 (1.173-4.433)[a], 2.249 (1.175-4.306)[b]		NS		0.013	2.259 (1.184-4.310)	
1	**rs513555**	***SPRR2G***	0.002	1.580 (1.17-2.137)[b], 1.338 (1.093-1.64)[c]	0.001	1.695 (1.244-2.309)[b], 1.323 (1.073-1.632)[c]		0.001	1.277 (1.113-1.468)	0.001	1.396 (1.145-1.703)	
1	**rs7542294**	***CHI3L1***	0.015	1.373 (1.024-1.841)[b], 1.348 (1.084-1.676)[c]	0.015	1.381 (1.103-1.729)[c]		0.012	1.198 (1.042-1.378)	0.004	1.369 (1.106-1.694)	
1	**rs7515776**	***CHI3L1***	0.039	1.292 (1.061-1.575)[c]	0.021	1.333 (1.087-1.635)[c]		NS		0.011	1.285 (1.058-1.561)	
1	**rs6691378**	***CHI3L1***	NS		NS			0.048	1.163 (1.003-1.35)	0.025	1.258 (1.029-1.538)	
1	**rs946261**	***CHI3L1***	0.015	0.650 (0.471-0.898)[a], 0.632 (0.458-0.873)[b]	0.022	0.653 (0.467-0.914)[a], 0.629 (0.45-0.879)[b]		0.024	0.849 (0.737-0.978)	0.006	0.641 (0.466-0.882)	
2	rs1020396	*ZAP70*	0.033	1.253 (1.03-1.526)[c]	NS			NS		NS		
3	**rs9877854**	***CD80***	0.022	2.481 (1.288-4.783)[a], 2.327 (1.231-4.401)[b]	0.017	0.469 (0.243-0.905)[b], 0.439 (0.223-0.864)[c]		NS		0.021	0.461 (0.239-0.888)	
3	**rs6778945**	***CD80***	0.013	2.682 (1.369-5.256)[a], 2.467 (1.289-4.726)[b]	0.03	2.556 (1.277-5.119)[a], 2.347 (1.20-4.591)[b]		NS		0.011	2.393 (1.226-4.671)	
3	rs1060679	*CLDN1*	NS		NS			0.034	0.815 (0.676-0.984)	NS		
3	rs3774032	*CLDN1*	NS		NS			0.033	0.744 (0.568-0.976)	NS		
3	**rs6776378**	***CLDN1***	0.006	0.474 (0.261-0.864)[b], 0.781 (0.63-0.969)[c]	0.026	0.505 (0.274-0.932)[b]		0.001	0.743 (0.62-0.893)	0.022	0.78 (0.630-0.965)	
5	**rs153449**	***GM2A***	0.020	1.336 (1.084-1.648)[c]	0.026	0.753 (0.607-0.935)[a]		NS		0.031	0.798 (0.650-0.979)	
6	rs9383692	*OPRM1*	0.043	0.720 (0.553-0.94)[b]	NS			0.016	0.848 (0.742-0.969)	NS		
6	**rs9479769**	***OPRM1***	0.043	0.659 (0.468-0.929)[a], 0.663 (0.472-0.933)[b]	0.036	1.514 (1.065-2.153)[b], 1.589 (1.115-2.264)[c]		NS		0.011	1.55 (1.105-2.173)	
6	rs790258	*OPRM1*	0.048	2.339 (1.043-5.244)[b]	NS			0.032	1.252 (1.023-1.533)	NS		
8	**rs420610**	***ASAH1***	0.012	1.430 (1.13-1.812)[a], 1.309 (1.008-1.702)[b]	0.019	1.42 (1.112-1.813)[a]		NS		0.007	1.367 (1.088-1.717)	
8	rs7460992	*ASAH1*	NS		NS			0.050	1.146 (1.002-1.312)	NS		
9	**rs4135188**	***TXN***	0.015	0.737 (0.6-0.907)[c]	0.014	0.729 (0.589-0.902)[c]		NS		0.006	0.756 (0.619-0.924)	
9	**rs10980288**	***TXN***	NS		NS			0.033	0.735 (0.557-0.972)	0.045	0.734 (0.542-0.993)	
9	rs4135168	*TXN*	0.043	9.703 (1.147-82.072)[a], 8.805 (1.057-73.327)[b]	NS			NS		NS		
9	**rs13284515**	***LCN1***	0.016	1.382 (1.097-1.742)[c]	0.019	1.389 (1.094-1.763)[c]		0.041	1.245 (1.012-1.533)	0.013	1.344 (1.064-1.698)	

Table 3-16: Significant tag SNPs re-analysed after controlling for gender and age group.

Chr	SNP	Gene	General P value	Unadjusted OR (95% CI)[a,b,c]	P value	Adjusted for gender and age grp OR (95% CI)[a,b,c]	Multiplicative P value	Unadjusted OR (95% CI)	P value	Adj for gender and age grp OR (95% CI)
11	**rs353625**	***CD44***	NS		NS		0.050	0.842 (0.712-0.998)	0.020	1.895 (1.531-2.344)
11	**rs3794116**	***CD44***	0.010	3.161 (1.075-9.296)[b], 1.308 (1.031-1.66)[c]	0.014	3.367 (1.11-10.21)[b], 1.296 (1.014-1.655)[c]	0.005	1.363 (1.099-1.692)	0.015	1.348 (1.060-1.715)
11	**rs353632**	***CD44***	0.030	1.319 (1.056-1.648)[c]	NS		0.011	1.295 (1.064-1.578)	0.019	1.307 (1.044-1.635)
11	**rs16927042**	***CD44***	0.011	1.388 (1.095-1.761)[c]	0.020	1.345 (1.052-1.718)[c]	0.003	1.380 (1.121-1.701)	0.008	1.38 (1.089-1.749)
11	**rs7934770**	***CD44***	0.021	0.749 (0.595-0.944)[c]	0.019	1.343 (1.058-1.705)[a], 0.453 (0.208-0.990)[c]	NS		0.042	1.271 (1.009-1.601)
11	**rs713330**	***CD44***	0.031	0.689 (0.521-0.914)[c]	NS		0.017	0.726 (0.56-0.943)	0.019	1.405 (1.056-1.867)
11	**rs3825072**	***CAPN1***	0.011	0.665 (0.451-0.982)[b], 0.770 (0.632-0.939)[c]	0.008	0.634 (0.424-0.948)[b], 0.765 (0.624-0.938)[c]	0.003	0.794 (0.682-0.925)	0.003	0.744 (0.612-0.904)
11	rs17146724	*CAPN1*	NS		NS		0.040	1.177 (1.009-1.373)	NS	
11	**rs2271448**	***CAPN1***	0.002	0.526 (0.324-0.857)[b], 0.724 (0.58-0.906)[c]	0.002	0.499 (0.302-0.825)[b], 0.73 (0.58-0.919)[c]	0.001	0.744 (0.626-0.884)	0.002	0.697 (0.558-0.871)
11	**rs17134961**	***C11orf30***	0.019	1.323 (1.085-1.616)[c]	0.035	0.77 (0.626-0.946)[a]	NS		0.030	0.805 (0.661-0.980)
12	rs11170124	*KRT6B*	NS		NS		0.028	0.676 (0.477-0.959)	NS	
12	**rs687751**	***KRT5***	0.029	0.741 (0.595-0.924)[c]	0.026	1.364 (1.087-1.711)[a]	NS		0.009	1.326 (1.073-1.639)
12	**rs607860**	***KRT5***	NS		NS		0.049	0.868 (0.757-0.998)	0.017	0.787 (0.647-0.958)
14	**rs2765901**	***RNASE7***	0.004	0.609 (0.445-0.834)[b], 0.804 (0.659-0.983)[c]	0.021	0.652 (0.472-0.902)[b]	0.001	0.789 (0.686-0.908)	0.017	0.787 (0.647-0.958)
14	**rs1243469**	***RNASE7***	0.023	1.322 (1.062-1.646)[c]	NS		0.007	1.300 (1.078-1.569)	0.017	0.766 (0.616-0.954)
14	rs17094917	*SERPINA12*	NS		NS		0.026	0.851 (0.739-0.98)	NS	
15	**rs894785**	***CTSH***	0.022	1.323 (1.009-1.735)[b], 1.382 (1.093-1.749)[c]	0.025	0.728 (0.571-0.928)[a], 0.725 (0.548-0.958)[b]	0.049	1.145 (1.002-1.309)	0.007	0.727 (0.578-0.915)
15	**rs8034542**	***CTSH***	0.017	0.750 (0.614-0.916)[c]	0.040	1.302 (1.060-1.599)[a]	0.016	0.830 (0.714-0.965)	0.013	1.282 (1.055-1.557)
15	**rs10400902**	***CTSH***	NS		NS		0.032	1.157 (1.013-1.322)	0.031	0.783 (0.626-0.978)
16	**rs2734741**	***PPL***	0.022	1.682 (1.16-2.441)[a], 1.505 (1.045-2.169)[b]	0.049	0.62 (0.422-0.91)[c]	NS		0.026	0.659 (0.457-0.951)
16	**rs2908662**	***PPL***	0.002	0.605 (0.455-0.805)[a], 0.705 (0.525-0.948)[b]	0.006	0.618 (0.461-0.830)[a], 0.71 (0.524-0.963)[b]	NS		0.003	0.657 (0.496-0.868)
16	**rs2451**	***PPL***	0.002	0.552 (0.393-0.777)[a], 0.601 (0.428-0.846)[b]	0.003	0.546 (0.384-0.776)[a], 0.602 (0.423-0.855)[b]	NS		0.001	0.574 (0.410-0.803)
18	**rs8089682**	***DSC3***	NS		NS		0.035	1.232 (1.019-1.491)	0.041	1.262 (1.009-1.578)
18	rs9304098	*DSG2*	NS		NS		0.021	1.178 (1.027-1.354)	NS	
18	**rs2290129**	***DSG2***	0.008	0.695 (0.537-0.900)[a], 1.277 (1.025-1.593)[c]	0.004	0.776 (0.618-0.974)[a], 1.509 (1.155-1.971)[c]	NS		0.013	1.384 (1.072-1.788)
19	**rs10426**	***KLK10***	0.001	1.529 (1.221-1.915)[c]	0.001	1.541 (1.222-1.942)[c]	0.001	1.383 (1.135-1.686)	<0.001	1.497 (1.194-1.876)
19	**rs10425377**	***KLK10***	NS		NS		0.029	1.182 (1.019-1.373)	0.029	1.242 (1.022-1.509)
19	**rs2569488**	***KLK14***	0.037	0.724 (0.563-0.932)[c]	0.037	1.405 (1.083-1.824)[a]	0.033	0.781 (0.624-0.98)	0.015	1.369 (1.062-1.763)
19	**rs6509518**	***KLK14***	NS		NS		0.037	0.814 (0.672-0.988)	0.045	1.255 (1.005-1.567)
19	**rs10500304**	***KLK14***	0.021	0.711 (0.553-0.916)[c]	0.024	1.42 (1.093-1.844)[a]	0.043	0.789 (0.63-0.99)	0.017	1.363 (1.058-1.755)
20	**rs742035**	***CTSA***	0.013	1.359 (1.052-1.758)[c]	0.021	1.379 (1.059-1.796)[c]	0.008	1.389 (1.093-1.765)	0.010	1.409 (1.085-1.831)

Odds ratio (OR) and 95% confidence interval (CI) at the aa vs aA [a], aa vs AA [b] and aA vs AA [c], NS- not significant, Adj- adjusted

3.4 Discussion

The validation study helped identify 27 of the 145 tested genes which are associated with AD in the ethnic Chinese population of Singapore. The 27 genes found significant contain a few known candidates whereas a few are novel.

3.4.1 Genes associated directly with barrier dysfunction

The gene with the maximum tags being independently associated with AD is carbonic anhydrase VI (*CA6*), a novel candidate identified and validated to be associated with AD. Of the 8 signals, minor alleles of 2 SNPs (rs1409147 and rs942969) are protective, whereas for the other 6 SNPs, the minor allele has a risk-conferring effect. 5 SNPs of the 8 have an allelic effect, whereas 3 SNPs (rs2274327, rs2274328 and rs1576766) have a genotypic effect with the heterozygous being risk-conferring. *CA6* is hypothesized to play a role in maintaining the skin pH thereby affecting the processes regulating desquamation and barrier integrity.

Filaggrin 2 (*FLG2*), a gene found on the epidermal differentiation complex of chromosome 1 was validated to be associated in a recessive model with the homozygous minor having a risk effect as compared to the heterozygous and homozygous major. Different variants in African-American population have been associated with persistent AD (Margolis *et al.*, 2013). However, *FLG2* was not found as a marker for chronicity when compared to non-chronic cases (Table 3.14) and when comparing chronic cases to controls. It was however, associated as a marker for non-chronic versus control with similar odds ratio of 2.6 (data not shown). As *FLG2* shares similar function and localisation with

FLG, it's hypothesized to be involved in a synergistic mechanism with *FLG* to maintain the barrier (Wu *et al.*, 2009).

Another gene from the epidermal differentiation complex of 1q21 repeatedly associated with AD is a structural protein, small proline rich protein 2G (*SPRR2G*). SPRR2G is from a group of proteins whose regulation in different environmental conditions provides adequate barrier function (Cabral *et al.*, 2001). Through mice studies, it has been shown that oestrogen is an important regulator of the members of the SPRR2 family (Hong *et al.*, 2004). In our study, when controlling for the effect of gender and age group, we do observe a slight increase in odds ratio from 1.581 [case-control association (Table 3.10)] to 1.695 [association after controlling for gender and age (Table 3.16)] for *SPRR2G*.

Another gene associated with AD is chitinase 3 like protein, *CHI3L1*, also known as YKL-40. It has been associated with asthma as its circulating levels have been correlated with the thickness of the epithelial basement membrane (Chupp *et al.*, 2007). Asthma like AD, has an epithelial barrier dysfunctional hypothesis of the lungs allowing for increased access of the allergens, pathogens and toxins to the airway tissues (Holgate *et al.*, 2009). Based on our data, odds of developing AD is 1.57 and 1.61 (unadjusted and adjusted OR) times higher with asthma (Table 2.13). *CHI3L1* is a known candidate for asthma, but a novel candidate for atopic dermatitis showing 4 independent signals of association, 3 of which being risk conferring and one having a protective effect.

Claudin-1 (*CLDN1*) is associated through 3 tag SNPs. Minor alleles of all three SNPs are protective in a dominant manner. Two of these SNPs rs3774032 and rs6776378 are also protective with the same strength of OR when compared with SPT positives. Claudin-1 is a tight junction protein found to be present in higher levels in lesional AD skin as compared to healthy controls as compared to non-lesional skin (De Benedetto *et al.*, 2008) suggesting a reduction in the baseline levels of CLDN1. As CLDN1 is a tight junction protein, a reduction in its levels would result in a less intact skin barrier resulting to greater ingress of allergens and pathogens manifesting in AD. SNPs from CLDN1 were later shown to be associated with AD in two North American populations (De Benedetto *et al.*, 2011). SNPs from *CLDN1* have also be associated with contact sensitisation to nickel and fragrances (Ross-Hansen *et al.*, 2013) indicating its vital role in maintaining the skin barrier.

Junction proteins like CLDN1 are important proteins in maintaining cell cohesion and thus intactness. Other members such as desmogleins and desmocollins forming the desmosomes are also associated with AD. Periplakin (PPL), a component of the desmosome, links the desmosome internally to the keratins and externally to the cornified envelope. Desmogleins, besides targets for the human proteases are also cleaved by Staphylococcal toxins (Hanakawa *et al.*, 2002). Staphylococcus infection is present in chronic AD, further supporting the observation of DSG2 as a marker for chronicity (Table 3.14). Interaction with DSG2 was also found to trigger epithelial to mesenchymal transition involving loss of intercellular junctions (Wang *et al.*, 2011). Also, DSG2 has been shown to be present in lower levels in AD skin as compared to

controls (Saaf *et al.*, 2008). SNPs in *DSG2* could either be attributing to the reduced expression thereby leaving the epidermis open for pathogens and allergens or the SNPs could be making the protein more sensitive to its ligands or pathogenic proteases.

Desmocollin 3 (*DSC3*) associated with AD is found in basal layers of the epidermis. It is shown to counteract the effect of excessive serine proteases by upregulating its own expression (Hachem *et al.*, 2006). Also auto-antibodies against DSC3 are found in another barrier dysfunction disease called pemphigus vuglaris (Kalantari-Dehaghi *et al.*, 2013). Reduction of DSC3 expression either through the SNPs of making it a better antigen for auto-antibodies predisposes an individual to a compromised barrier leading to AD phenotype.

Periplakin is part of the scaffold on which the cornified envelope (barrier property) of the outermost epidermal layer develops. Knock-out mice models have shown delayed embryonic barrier formation with reduced lipid content and mechanical strength (Sevilla *et al.*, 2007). Defective expression of PPL has also been observed in AD (Guttman-Yassky *et al.*, 2011). Studies in combination with our data strongly associate *PPL* with AD pathogenesis.

GalNAc4(Neu5Ac3)Gal4Glc-ceramide ganglioside activator (*GM2A*) is found in the lamellar bodies of the epidermis and is a glycolipid transport protein (Raymond *et al.*, 2008). Mice studies have shown the acquisition of epidermal barrier by transcription factor Klf and corticosteroids involves the up-regulation of GM2A by 22 fold amongst other downstream products (Patel *et al.*, 2006). These results clearly highlight the plausible role *GM2A* plays in

barrier maintenance and based on our data, its association to AD makes it a novel candidate for AD.

N-acylsphingosine amidohydrolase (acid ceramidase) 1 (*ASAH1*) is associated with AD through 2 SNPs. rs420610 has a strong risk effect in a recessive fashion with OR as 1.4 and another SNP rs7460992 confers a marginal risk at the allelic level (OR 1.14). The mature protein catalyses the synthesis and break-down of ceramides into fatty acids and sphingosine which help in retaining moisture and act as antimicrobial agents respectively. In an elegant study, Arikawa *et al* have shown the association of reduced sphingosine to increased *Staphylococcus aureus* colonisation from involved and uninvolved lesions of AD patients versus healthy individuals. They also showed reduced acid ceramidase correlating with decreased sphingosine and ceramide (Arikawa *et al.*, 2002). Thus suggesting the reduced levels of ASAH1 contributes to an impaired antimicrobial barrier resulting in AD.

Calpain1 is associated with AD via 4 tag SNPs 1 of them is a strong marker for chronicity. 3 of these SNPs remain significant when regressing for gender and age. Calpain1, secreted by the lamellar body (Raymond *et al.*, 2008), is involved in the processing of the profilaggrin in differentiating keratinocytes (Yamazaki *et al.*, 1997). Filaggrin is the most widely studied and well replicated gene associated with AD. If the associated SNPs are involved in a change in activity of calpain then it would lead to aberration in the processing of filaggrin leading to impaired skin barrier.

Keratin 5 and 6B are type II cytokeratins co-expressed during differentiation of stratified epithelium. Keratins form the key structural component and an

aberration in their synthesis or structure may in-turn affect the integrity of skin. Previous studies have observed over-expression of KRT6A, 6B in chronic AD lesions highlighting their role in pathogenesis (Jarzab *et al.*, 2010). On regressing with gender and age group, *KRT6B* loses significance whereas *KRT5* remains significant (Table 3.16).

Ribonuclease, RNase7, is a broad spectrum antimicrobial peptide found on the skin. It is a part of the innate immunity by forming an antimicrobial barrier thereby preventing pathogens from colonising. RNAse7 is found in lower levels of unaffected AD skin as compared to control's (Zanger *et al.*, 2009) indicating a broad antimicrobial barrier failure prior to disease manifestation. RNase7, like other antimicrobial peptides are inducible upon injury and were found to be produced in significantly higher levels in AD as compared to controls (Harder *et al.*, 2010) indicating regulation of the skin's antimicrobial environment. *RNASE7* is associated by 2 independent signals, one having protective and the other having a risk-conferring association.

Proteases and protease inhibitors regulate the activation of different important enzymes and help regulate the differentiation and desquamation procedures. As members from these two group of proteins regulate homeostasis of the barrier, they become obvious choice as candidates for dysregulation. Like in other studies, we have found proteases such as kallikreins and cathepsins, and protease inhibitor from the serine protease inhibitor (SERPIN) family associated with AD.

The usual serine protease candidates are KLK5 and KLK7, however, we have detected KLK10 and KLK14 to be associated with AD in our population.

KLK10 and KLK14 have been shown to be up-regulated in SC of AD patients (Komatsu *et al.*, 2007). However, it remains to be ascertained if the levels are associated with the significant genotypes in our population. KLK14 has been shown to activate the pro-protein forms of KLK5, KLK1 and KLK11 (Emami and Diamandis, 2008) and to cleave desmogleins (junction proteins) at a greater rate than KLK5 (Brattsand *et al.*, 2005). KLK14 is also correlated with increased trans-epidermal water loss and reduced hydration (Stefansson *et al.*, 2008), highlighting it's critical role in maintaining the integrity of the permeability barrier. All 3 SNPs from *KLK14* have a protective role whereas the 2 SNPs from *KLK10* have a risk-conferring role with one of them having an OR of 1.5. However, after controlling for the effects of demographics, 1 SNP from *KLK14* and both SNPs from *KLK10* remain significant.

We tested all members of the cathepsin family of cysteine protease and found cathepsin A (*CTSA*) and cathepsin H (*CTSH*) to be associated in our population. Cathepsin S over-expressing transgenic mice induced scratching behaviour via proteinase-activated receptor-2 (PAR-2) showing chronic AD symptoms (Kim *et al.*, 2012). Low levels of cathepsin E were observed in AD subjects and mice models lacking this protease spontaneously developed AD lesions suggesting its role in pathogenesis (Tsukuba *et al.*, 2003). Reduction in cathepsin G positive cells is also shown to be associated with clinical improvements (Breuckmann *et al.*, 2002) linking reduced protease to skin improvement. There is no direct reported association of *CTSA* and *CTSH* to AD, but based on the links to other members of the same family, these 2 novel validated candidates provide a direction for future research.

One SNP from *SERPINA12* was found associated in case versus control analysis however, 4 with chronicity. As mentioned above the importance of serine protease in regulating homeostasis of the epidermis, their regulation by inhibitors also becomes crucial for the same. Mice studies have shown the acquisition of epidermal barrier by transcription factor Klf and corticosteroids involves the upregulation of SERPINA12 by more than 10 fold amongst other downstream products (Patel *et al.*, 2006) highlighting its role in barrier acquisition. Dyregulation of SERPINA12 is also associated with psoriasis, another barrirer dysfunction disease (Gudjonsson *et al.*, 2010).

3.4.2 Genes associated indirectly with barrier dysfunction

Heterogeneity of ZAP70, zeta-chain of T-cell antigen receptor associated protein kinase deficiency has been reported to present atopic dermatitis in some patients (Turul *et al.*, 2009). Others have suggested deficiency of ZAP70 predisposes one to exaggerated inflammation and autoimmunity (Poliani *et al.*, 2013; Roifman *et al.*, 2010). Besides, barrier dysfunction and atopy, autoimmunity is also a suggested mechanism of pathogenesis. Altered T- cell receptor signalling is also a known cause of allergic diseases (Datta and Milner, 2011). This could explain the loss of significance when testing AD cases against SPT positives (Table 3.12), besides the lack of power to detect association with such low numbers. From Tables 3.14 *ZAP70* appears to be a marker for non-chronic AD.

Cluster of differentiation/designation (CD) 80 is strongly associated with AD in a recessive manner with OR ranging from 2.3-2.5 for the 2 SNPs. The OR slightly increases when tested against SPT (Table 3.12) for rs6778945 and a third SNP gets highlighted to be significant in a protective manner indicating

independent and opposing effects of the minor alleles. CD80 binds to CD28/ CTLA4 activating Th2 subset of T-helper cells mediating allergic inflammation (Lanier *et al.,* 1995). CD80, CD86 mapped to 3q21 has been reported as a major susceptibility locus for AD and allergic sensitisation through genome wide linkage study (Lee *et al.,* 2000). CD80 has been found to be associated with AD at meta-analysis (Cao *et al.,* 2009). Others have shown that Langerhans cells and inflammatory dendritic epidermal cells show increased expression of CD80 and CD86 in AD skin indicating its role in pathogenesis of AD (Schuller *et al.*, 2001).

Opioid receptor, μ1 (OPRM1) is expressed on skin epidermal keratinocytes and melanocytes. They are responsible for regulating pain and itch sensations particularly in chronic dermatitis. They are also involved in skin differentiation , wound healing and scar formation. Signal transduced by μ opioid receptor is known to stimulate itch (Bigliardi *et al.,* 2009). Enhanced itch sensation further leads to the scratch-itch cycle promoting barrier dysfunction and thus AD. OPRM1 is associated through 3 tag SNPs in a recessive manner, 2 of the SNPs having a protective role whereas one has an OR of 2.2 (Table 3.10). When analysed in chronic vs non chronic cases 4 SNPs turn out significant with 3 being markers for non-chronic AD and 1 for chronic AD (Table 3.14) .

Thioredoxin (*TXN*) is associated with AD through 3 tag SNPs 2 of them protective in a dominant manner and the third SNP is risk-conferring in a recessive fashion. The protective association gets stronger when controlling for gender and age however, the risk-conferring SNP loses significance. Thioredoxin is a cross reactive allergen from fungus known to contribute to

asthma and AD. Studies have shown auto and cross reactivity to the human form by IgE from patients (Glaser *et al.*, 2008). Thioredoxin may thus be involved in the pathogenesis of AD through atopy and auto-allergy.

Lipocalin1 (*LCN1*) is associated through 1 SNP with OR 1.3 which remains significant after controlling for gender and age. Lipocalins are proteins that bind to various hydrophobic ligands and transport them extracellularly. They are inducible under stress and infection. Most of the mammalian allergens belong to the family of lipocalins (Chapman and Wood, 2001). The similarity of human and allergen source lipocalin may be involved in the manifestation of atopy (Virtanen *et al.*, 1999), thus leading to AD. Also, LCN2, another member of the same family has been shown to be downregulated in lesional AD skin along with other innate defense proteins (Guttman-Yassky *et al.*, 2008).

CD44 was associated with AD through 6 tag SNPs and all remained significant after controlling for demographics (Tables 3.10 and 3.16). One of the SNPs was found to be a marker for chronicity (Table 3.14). CD44 is implicated in a variety of functions ranging from cell-adhesion, migration, lymphocyte activation and homing to tumor metastasis. Diversity in function for CD44 is attributed to its complex alternative splicing. *Ex vivo* experiments on eosinophils from AD patients showed a significant higher expression of CD44 as compared to healthy controls' (Ogawa *et al.*, 2003). Knock-out mice experiments have shown the importance of CD44 in committing to leucocyte extravasation at the site of allergen challenge (Gonda *et al.*, 2005). Also, disruption of the epidermal permeability barrier increases CD44 expression

(Man *et al.*, 2009). Thus, CD44 is associated with AD via the atopy and barrier dysfunction hypothesis.

3.4.3 Miscellaneous association

CAP1 is a serendipitous association as the gene of earlier interest was PRSS8 which is also known as *CAP1* from chromosome 16, which has been associated with barrier dysfunction (Frateschi *et al.*, 2011; Leyvraz *et al.*, 2005). However, while tagging, we tagged another gene adenylate cyclase-associated protein 1 abbreviated as *CAP1* from chromosome 1 in addition to *PRSS8*. *PRSS8* wasn't significant in the screening but *CAP1* was and thus was put through validation and it remained significant with 2 tags being associated at the genotypic level. Thus, this is also a novel association to AD.

C11orf30 was tagged at 5% MAF at the screening round yielding 11 tags representing the gene (2 kb upstream and 1 kb downstream) out of which 1 tag was significant. Due to limitation in the number of SNPs on the chip, instead of re-genotyping the 11 tags on the larger validation population, only the significant SNP was genotyped and it was validated. A SNP 38kb downstream of *C11orf30* was reported through a GWAS for AD (Esparza-Gordillo *et al.*, 2009) and for Crohn's disease which is another barrier dysfunction disease (Barrett *et al.*, 2008). This gene codes for a nuclear protein which inactivates *BRCA2* and is known to localise at DNA repair sites. It is found to be up-regulated in epithelial tissue cancers of breast and ovaries (Hughes-Davies *et al.*, 2003).

3.5 Conclusion

The study yielded 27 candidate genes associated with AD **six** were novel. Most of the candidates can be attributed to AD through barrier dysfunction hypothesis. The ones not directly associated with maintaining the barrier such as candidates of atopy, sensitivity to excessive itch, auto-allergy can also indirectly lead to a compromised barrier.

This study has found associations to genes which were hypothesized based on their function but never actually shown genetically. The vast types of genes associated such as proteases, tight junction proteins, antimicrobial peptides, CD markers and genes overlapping with asthma shows the spectrum of mechanisms involved in AD's pathogenesis given a geographical and genetic background.

Chapter 4

Functional characterisation of *CA6*

4.1 Introduction

The results from the previous chapter point towards a multiple set of genes associated with AD. The genes implicated are either a direct risk of barrier dysfunction or an indirect risk leading to barrier dysfunction. Among the set of 27 genes found associated, carbonic anhydrase VI (*CA6*) has the maximum number of tags associated with AD and is a novel candidate. Thus, *CA6* was decided to be taken forward to evaluate functionally. Although the remaining 26 genes are also significant and work on them is in progress, *CA6* being a novel candidate and having the maximum number of associated SNPs was chosen for further characterisation.

4.1.1 Exome and alternative splicing

Excess of proteome when compared to the genome explains alternative splicing as a mechanism for expanding information from a limited source (Nilsen and Graveley, 2010). In recent years, alternative splicing has been estimated to be present in over 95% of the genes containing multiple exons (Blencowe, 2006). As such, differential splicing patterns serve as a key area for gene expression (Zhou *et al.*, 2013). *CA6* like many genes is known to undergo alternative splicing. To check whether alternative splicing occurring in *CA6* is associated with any of the significant SNPs, exon array study was conducted in an *ex vivo* manner. Differential splicing would provide

information on the exons being spliced out and whether selective patterns of splicing are predisposed by the associated SNPs.

4.1.2 Epigenetic methylation

Epigenetic regulation by methylation has been shown to have a regulatory effect (Lorincz *et al.*, 2004) to processes such as gene expression (Phillips, 2008) and splicing (Maunakea *et al.*, 2013). To better understand our gene expression and alternative splicing results, we undertook a methylation study and analysed the patterns with the associated SNPs. As the downstream of genetic and epigenetic regulations finally affects the protein levels, stratum corneum collected from the participants (Dreher *et al.*, 2005) was tested for CA6 levels via immunoblotting.

4.1.3 pH

The gene of interest - *CA6*, which although doesn't have a known function on the skin, is implicated to regulate pH (Khalifah, 1971a) primarily in the alimentary canal (Parkkila *et al.*, 1994), and in milk (Karhumaa *et al.*, 2001). The outer layers of stratum corneum is long known to have an acidic pH (Hetjss, 1892). Studies have shown a biphasic curve wherein the pH approaches neutrality in the lower layers of SC and then reaching a more acidic pH at the interface of the SC and stratum granulosum to then shift to neutrality again (Fluhr and Elias, 2002). Increased pH, considered a cause (Cork *et al.*, 2009) and an effect of barrier dysfunction (Ohman and Vahlquist, 1994), is an important physiological characteristic worth noting in AD. As a result, we measured skin pH on the re-called population of cases and controls and their readings were analysed in association to the significant SNPs.

4.1.4 Trans-epidermal water loss

Another hallmark of a dysfunctional barrier and AD is increased trans-cutaneous water loss associated with dry skin (Fluhr *et al.*, 2006). As explained in Chapter 1, skin with more physical gaps due to inherent aberrations or external sources will provide route for evaporative water loss. To measure this marker of a compromised barrier, trans-epidermal water loss (TEWL) was measured on re-called subjects and its association to the significant SNPs was tested.

4.1.5 *In vitro* characterisation

The different associations pointed towards functionality of the SNPs which requires further research. However, it is unlikely for a single SNP to have a strong enough effect to cause a disease. To cause a strong effect, the SNP in combination with other SNPs around it gathers the biological function to execute or modify a function. A combination of SNPs present on a chromosome which are passed down together are haplotypes. Thus, haplotype becomes the unit of functional inheritance within a gene. To evaluate the effect of the associated variants to gene function and ultimately to disease manifestation/ progression, we first found the haplotypes within our working population. As the significant SNP was a tag SNP, all SNPs in LD with the tag SNP were also considered significant and evaluated for a causal role. Haplotypes spanning significant SNPs were identified either through resequencing and/or by using HapMap data. Haplotypes prevalent at a frequency of 5% or higher in the population were considered for further functional characterisation.

4.2 Materials and method

4.2.1 Differential splicing and gene expression

To evaluate the differential spliced variants present *ex vivo*, samples were subjected to Affymetrix's Human Exon ST1.0 microarray. This array measures each exon by a probe giving higher probe density and better coverage than other conventional arrays available. Besides measuring differential exon levels resulting from alternative splicing, the chip also allows for measuring differential gene expression (Okoniewski *et al.*, 2007). RNA extracted was sent to Origen Laboratory where they performed the quality checks and ran the microarrays and raw data was provided to the lab.

4.2.2 Regulation by differentially methylated regions

To further evaluate the next layer of regulation involved, we undertook an epigenetic study to identify genome wide differentially methylated regions. Stock genomic DNA was subjected to Infinium HumanMethylation450 BeadChip (Illumina) and results were analysed for differences between methylation sites between cases and controls and between genotypes irrespective of disease status.

4.2.3 Stratum corneum collection and protein quantification

The SC was collected from participants at the inner flexure of the elbows by tape stripping. Tape discs of 22mm diameter (CuDerm) were stamped at and around the folds within a defined area. The protein was extracted by freeze-thawing and sonication and later subjected to immunoblotting for measuring relative levels of CA6. Analyses was done to check for differences between case-control and between genotypes.

4.2.4 TEWL and pH measurements

Trans-epidermal water loss and skin pH were measured on a recalled population. TEWL was measured with the help of Tewameter TM 300 MDD (Courage + Khazaka electronic GmbH) and pH was measured using the same company's pH probe 905. The recalled participants' skin conditions were equilibrated to the measuring environment's condition by letting them sit in the same room for 30-45 min before starting the measurements. TEWL and pH readings from both arms were taken. The participants were then tape stripped on both arms in a defined area. After tape stripping, TEWL and pH readings were taken again from both arms. The recalled population was called for two visits and the same procedure was conducted again on the 2nd visit. Differences between the genotypes of the significant tag SNPs were analysed against different TEWL and pH parameters.

4.2.5 Statistical analyses.

Excel (Microsoft) and SPSS (IBM) were used for the different analyses. For all the experiments, due to the limited sample size, non-parametric analyses such as Mann-Whitney-Wilcoxon and Kruskal Wallis methods were applied. As the epigenetic methylation study was a whole genome study, analysis was done with the parametric equivalent tests of analysis of variance (ANOVA). *P* value of less than 0.05 was considered significant.

4.2.6 Functional predictions

The significant SNPs and the SNPs in LD with the tag SNPs were evaluated for their putative functions through different databases. Depending on the region where the haplotypes spanned, functional predictions for various functions were done. Significant regions within the 5' untranslated region (5'

UTR) or upstream of the gene or in the first intron (Coulon *et al.*, 2010) were evaluated for promoter activities. Significant regions within the intron-exon boundaries were evaluated for their putative effect on splicing. Overall, functions of the SNPs were evaluated in SNPinfo (http://snpinfo.niehs.nih.gov/snpinfo/snpfunc/), F-SNP (http://compbio.cs.queensu.ca/F-SNP/). SNPs that lay in the upstream or first intron were evaluated for promoter activity through TFSearch (http://www.cbrc.jp/research/db/TFSEARCH.html) and ConSite (http:/www.phylofoot.org/consite). SNPs within the 300 bases of the intron-exon boundaries were predicted for their putative splicing functions through Human Splicing Finder (http://www.umd.be/HSF/). SNPs in 3'UTR were evaluated for miRNA based regulation through a miRNA database (http://www.mirbase.org/). Overall information on domains coded by different exons was sourced from Uniprot database (http://www.uniprot.org/).

4.2.7 Re-sequencing

2 kb upstream area of the gene was re-sequenced to find the haplotypes prevalent in the working population. Re-sequencing was performed by Sanger's dideoxy-nucleotide method (Sanger *et al.*, 1977). 40 samples randomly chosen from the cohort of Singapore Chinese samples were selected for sequencing to be able to identify SNPs with a frequency of 1.25% or more. Briefly, the area of interest spanning 2 kb upstream and 1kb into the start of the gene was amplified using a high fidelity polymerase such as Kapa HiFi (Kapa Biosystems) or Phusion (Thermo Scientific). This product was then re-sequenced using BigDye terminator kit (Applied Biosystems) spanning areas around significantly associated SNPs and around the intron-exon boundary.

SNPs confirmed by re-sequencing in forward and reverse direction were then identified by comparing to dbSNP (http://www.ncbi.nlm.nih.gov/projects/SNP/) to check for known and novel SNPs found through resequencing.

4.2.8 LD and haplotypes

SNPs found through resequencing, HapMap and Singapore Genome Variation Project (SGVP) (http://www.statgen.nus.edu.sg/~SGVP/) were tested for LD and haplotypes were obtained using Haploview (http://www.broadinstitute.org/scientificcommunity/science/programs/medical-and-population-genetics/haploview/haploview). Haplotypes with a MAF of less than 0.05 were excluded. Constructs for the haplotypes were either made in the lab through cloning procedures or purchased from Genscript.

4.2.9 Cloning

Clones for the promoter assay were prepared using routine cloning procedures. Primers containing restriction enzyme (RE) sites were used to amplify the different haplotypes. PCR product and the vector pGL4.10 (Promega) were digested with the set of XhoI-EcoRV and NheI-HindIII (New England Biolabs) REs and then ligated with T4 ligase. The ligated product was transformed in TOP10 cells (Invitrogen) and successfully transformed colonies were picked for plasmid extraction. Wizard Plus SV Miniprep DNA Purification System (Promega) was used to extract the plasmid DNA and checked for the correct sequence through sequencing. Colonies with correct sequence were then scaled up to grow in 200ml Luria Bertani broth and plasmid was extracted using PureLink HiPure Plasmid Filter Purification Kits

(Invitrogen). The sequence was again verified, quantified and used for downstream cell culture work.

4.2.10 Cell lines and transient transfection

Cell lines of mammalian origin were used to avoid the genetic variability associated with primary cells. Cell lines such as HaCat and HEK293T were used for the different in vitro assays. HaCat cells are normal keratinocytes which underwent spontaneous immortalisation to result in a nontumorigenic cell line (Boukamp *et al.*, 1988). HaCat were grown and maintained in Dulbecco's Modified Eagle Medium (DMEM, Sigma-Aldrich) supplemented with 10% foetal bovine serum (FBS, GIBCO). Another cell line namely HEK293T were also used for the *in vitro* experiments. Although not originated from skin, HEK293T are well known for their ease of transfectability and were used to counter cell-line specific bias. HEK293T were grown in RPMI-1640 (Sigma-Aldrich) supplemented with 10% FBS. All cell lines were grown in a greater than 80% humidity, 5% CO_2 environment at 37°C.

Cells were grown in 24 well plates until they reached more than 80% confluency. The plasmids containing haplotypes of interest were then transiently transfected using Lipofectamine 2000 (Invitrogen) according to manufacturer's instruction. The optimised ratio of DNA: Lipofectamine was 1:3 for the splicing experiment and 1:5 for the promoter experiment. Cells were harvested at 24 and/or 48hrs after transfection.

4.2.11 *In vitro* promoter activity assay

Promega's Dual Luciferase reporter system was used to measure promoter activity. The pGL4.10 vector contains a promoter-less fire-fly luciferase. The

test haplotypes were inserted in the multiple cloning site of this promoter which is upstream of the gene coding fire-fly luciferase enzyme. The recombinant pGL4.10 was co-transfected with pGL4.74 which has a HSV-TK promoter in front of renilla luciferase. The ratio of firefly luciferase to renilla luciferase used was 400:1. This was transfected in HaCat and HEK with Lipofectamine at a ratio of 1:5. Cells were harvested at 24hr for the upstream contruct and at 48 hrs for the intron 1 construct. The cells were lysed with the lysis buffer provided in kit and subsequent steps followed as per manufacturer's instructions. Luminescence was read in a Veritas Microplate Luminometer (Turner Biosystems). Luminescence was normalised by taking the ratio of firefly's luminescence over renilla luminescence readings and checked for differences between haplotypes. The experiment was conducted in triplicates and repeated a minimum of three times. Difference were checked with student's t-test.

4.2.12 *In vitro* differential splicing assay

The mini-gene assay was used to check for the differences in splicing patterns of the 2 constructs (Gaildrat *et al.*, 2010). As the constructs involved joining different exons with their 200-300 bases flanking introns, to be compressed into a mini-gene, the constructs were ordered rather than made in the lab. The constructs were ordered in the mammalian expression vector pcDNA3.1. Post 24 hrs of transfection, the cells were harvested and RNA extracted using E.Z.N.A.Total RNA kit (Omega). Equal amounts of RNA from the different haplotypes (400ng) was converted to cDNA using iScript cDNA Synthesis kit (Bio-Rad). To selectively amplify the regions of interest, primers in extreme exons were used to PCR the cDNA. Fluorescent dye FAM was run along in

the PCR to be incorporated by a low-fidelity Taq polymerase (Vivantis). The presence and relative abundance of the splice variants were analysed qualitatively by running a gel and quantitatively by running a capillary electrophoresis.

4.3 Results

CA6 is a large 29.2 kb gene on chromosome 1 and is made up of (at maximum) 9 exons. *CA6* had 8 tags associated with AD after the validation study. Following table 4.1 is a recap of the association of the significant SNPs from *CA6*. The risk allele and genotypes has been identified by the odds ratio and presented here in terms of A and B for ease of understanding.

SNP	P_Geno	OR (95% CI)	Geno*	Ref geno	P_allelic	OR	(95% CI)	Allele*
rs1409147	NS				0.043	0.829	(0.693-0.992)	B
rs942969	0.045	0.452 (0.22-0.93)[b]	BB	AA	0.038	0.825	(0.689-0.987)	B
rs2274327	0.009	1.361 (1.118-1.658)[c]	AB	BB	0.013	1.212	(1.043-1.409)	A
rs2274328	0.022	1.344 (1.086-1.664)[d]	AB	AA	0.043	1.150	(1.005-1.315)	B
rs11576766	0.015	1.265 (1.038-1.541)[d], 0.664 (0.469-0.939)[e]	AB	AA and BB	NS			
rs17032921	0.035	2.09 (1.123-3.889)[b]	AA	BB	0.026	1.236	(1.029-1.484)	A
rs7545200	NS				0.032	1.187	(1.017-1.386)	A
rs3765965	0.025	2.19 (1.064-4.507)[f]	BB	AA	0.012	1.285	(1.059-1.56)	B

a- AA vs AB, b- AA vs BB, c- AB vs BB, d- AB vs AA, e- BB vs AB, f- BB vs AA

* Risk variant, geno-genotype, Ref geno- Reference genotype

Table 4-1: Significant associations of *CA6*

CA6 being a large gene has multiple spliced variants. Aceview from NCBI shows presence of different reported spliced variants.

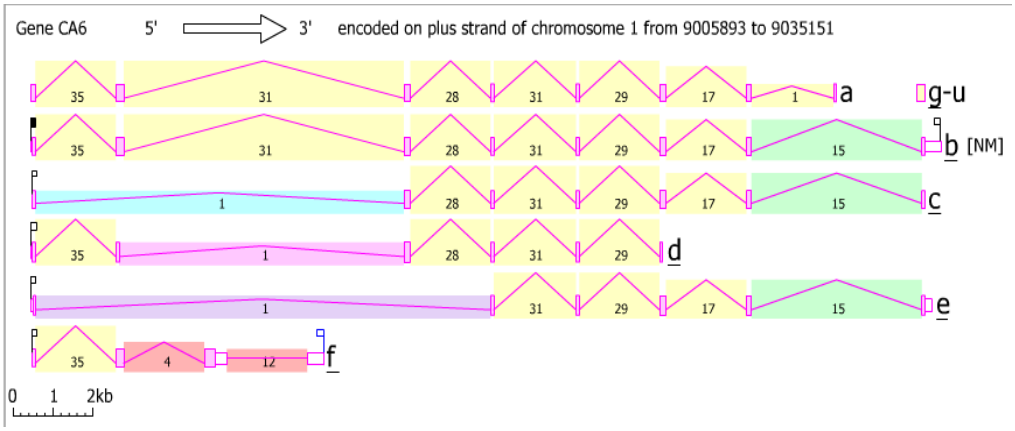

Figure 4-1: Different spliced forms of CA6
Different alternatively spliced variants of CA6 as reported in AceView (NCBI)

4.3.1 Exome array

To evaluate the presence and effect of the associated SNPs, we undertook a whole genome exome analysis on 20 samples as described in section 4.2.1. Data shown here is for the gene of interest- *CA6* alone. There were 16 probes in the Affymetrix Human exon array ST1.0 array to probe for the different exons of *CA6*. For obtaining the exon probe data, the exon scores were normalised with respect to the complete gene expression scores. Genotypes of the 8 significant tags were analysed with respect to the gene expression and probe specific data. The table below (Table 4.2) compiles only the significant associations (*P* values < 0.05) found between the 8 SNPs' genotypes with respect to gene expression and the 16 exon (probe) specific data.

SNP	Gene expression	Upstream 5'UTR p_2319138	E1 p_2319141	E9 p_2319155	3'UTR after E9 p_2319156
rs1409147					
AB vs BB		0.018			
rs2274328					
AA vs AB			0.049		
AB vs BB	0.028			0.019	
rs11576766					
AA vs AB	0.02				0.02
AB vs BB	0.01	0.019		0.01	

Table 4-2: Exon probes associated with the significant SNPs.

The table shows the region in the gene where the probe sets map to. Of the 8 SNPs analysed against the 16 probes, 3 SNPs showed association in 4 of the probe regions, and 2 SNPs showed association with gene expression.

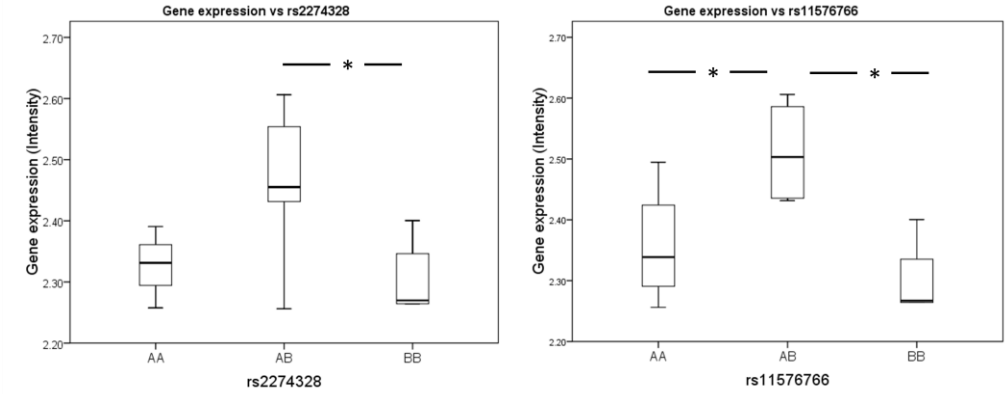

Figure 4-2: Gene expression data for associated SNPs.

Figure 4.2 shows the gene expression data with respect to the 2 significant genotypes. For rs2274328, gene expression was significantly different between the AB and BB genotypes. As shown in Table 4.1, AB genotype is the risk-conferring genotype indicating that the it leads to enhanced gene expression of *CA6*.

For rs11576766, AB genotype showed the significantly higher gene expression as compared to AA and AB. As shown in Table 4.1, AB genotype is the risk-conferring genotype indicating that the it also leads to enhanced gene expression of *CA6*, similar to rs2274328.

This suggests that overall cases have a higher expression of the *CA6* gene

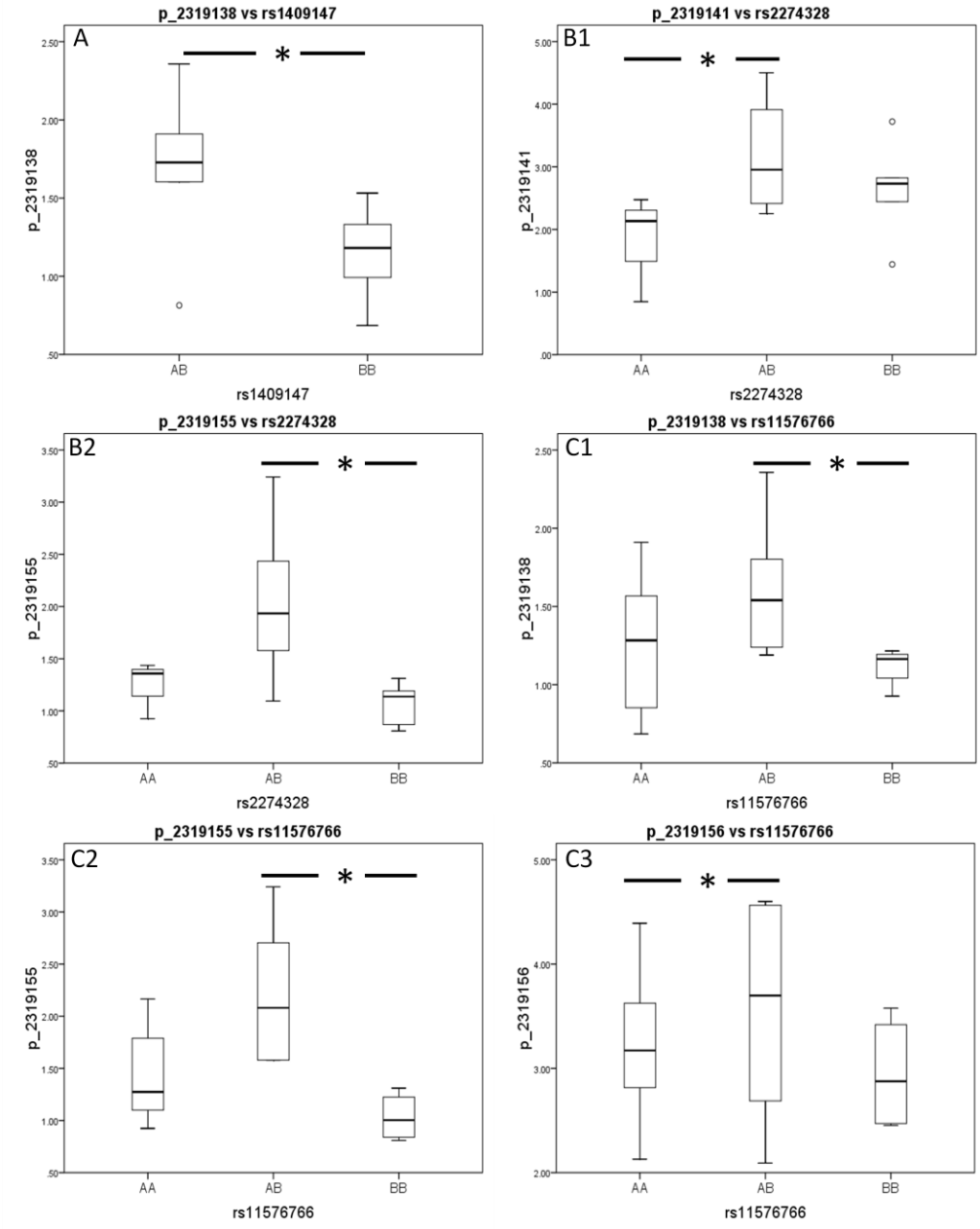

Figure 4-3: Exon-probe expression data for associated SNPs.

Exon probe expression for: p_2319138 vs rs1409147 (A), p_2319141 (B1) and p_2319155 (B2) vs rs2274328, and p_2319138 (C1), p_2319155 (C2) and p_2319156 (C3) vs rs11576766.

Figure 4.3 shows the exon probe expression data with respect to the significant SNPs as indicated in table 4.2. SNP rs1409147 found upstream of the gene was associated with the probe probing for exon 1 of the gene, suggesting the SNP affects the expression of exon 1. The risk-conferring allele as described in table 4.1 is B. The BB genotype is shown to have a lower expression of exon 1 as compared to the heterozygous state (Figure 4.3A).

Figure 4.3B shows the difference in exon expression data when compared with rs2274328. AB genotype which confers risk shows increased levels of exon expression for exon 1 (Figure 4.3B1) and exon 9 (Figure 4.3B2). This is in the same trend as the overall gene expression by the same SNP.

Figure 4.3C shows the exon expression data for rs11576766 with respect to 3 different probes. The 3 probes are present upstream of the gene, in exon 9 and in the 3'UTR after exon 9. The risk-conferring genotype AB, shows an increased trend for the regions following the same trend as gene expression in Figure 4.3.

To understand the mechanism by which the above SNPs affect gene expression and alternative splicing, we undertook a methylation study.

4.3.2 Methylation study

To evaluate the regulatory effect of the associated SNPs, a genome-wide epigenetic study to find differentially methylated regions was undertaken on 19 AD cases and 20 controls using Illumina's technology as described in

section 4.2.2. Data analysed is only for the 11 probes that spanned *CA6* and regions around it. Table 4.3 below contains the compiled results for only the significant probes (*P* value <0 .05) with respect to the genotypes. As the dataset contained a sizeable number of cases and controls, analysis was also carried out to check for differences in the methylation with respect to disease.

	After E1	Bet E5 and E6	Upstream of gene		Bet E6- E7
SNP	cg25919221	cg07601804	cg23738308	cg24052039	cg24885723
Case vs control				0.014	
rs942969					
AB vs BB	0.024				
rs2274327					
AB vs BB					0.024
rs2274328					
AA vs AB		0.041			
AA vs BB		0.016			
rs11576766					
AA vs AB		0.014	0.028		
AA vs BB		0.046			
rs17032921					
AB vs BB		0.005			
rs3765965					
AA vs AB		<0.001			

Table 4-3: Methylation probes associated with the significant SNPs.

The top row of the table gives information of the position of the probes. There was a difference in the methylation pattern of cg24052039 between cases and controls. And this probe maps to the upstream area of the gene, suggesting regulation of *CA6* via methylation affects disease outcome as shown in figure 4.4A below.

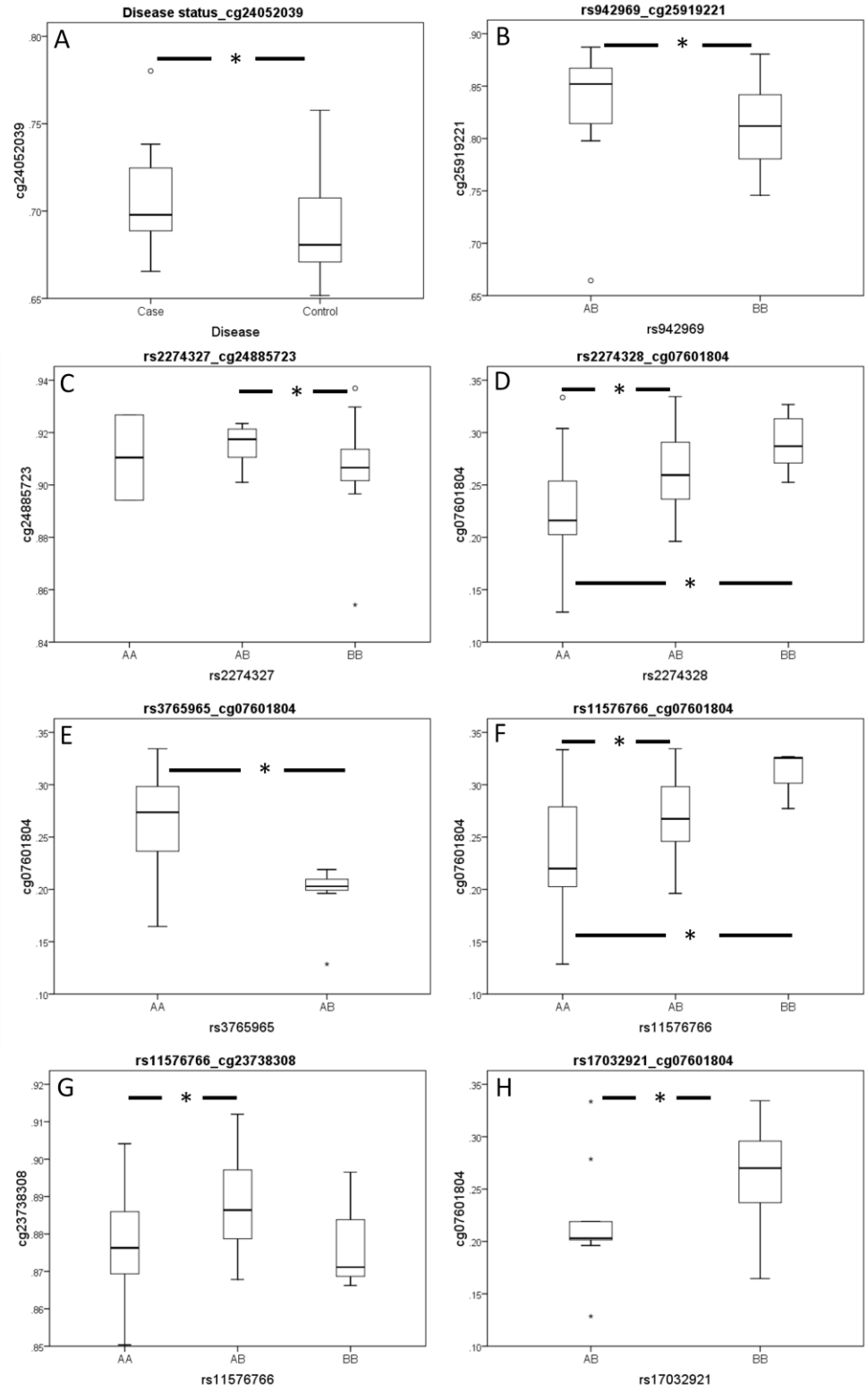

Figure 4-4: Methylation profile of associated SNPs and disease status.

Figure 4.4A shows that cases have a slightly higher methylation profile at the given upstream marker. As methylation in the upstream region of the genes

results in gene-silencing, these results indicate that cases having down-regulation leading to low expression of *CA6*.

The intron 1 SNP rs942969 showed decreased methylation in the risk-conferring genotype BB. The decreased methylation found after E1 could either result in increased expression of *CA6*, which is in line with the exome array or could affect alternative splicing (Figure 4.4B).

The SNP in rs2274327 shows increased methylation in AB risk-genotype (Figure 4.4C). The location of this probe is between E6 and E7 indicating a role in splicing. Similarly, rs2274328 showed increased methylation in AB (risk-conferring genotype) and highest in BB (risk-conferring allele) (Figure 4.4D). The probe lies between E6 and E7 and is known to affect the area between the introns (Illumina information).

SNP rs3765965 shows a stark difference in the methylation patterns between AA and AB. The samples put through the methylation did not have a BB for rs3765965 as B is the minor allele. From the graph (Figure 4.4E), it can be interpreted that BB might have a reduced expression as the presence of just one B in the heterozygous state caused a large difference. The probe associated with this SNP lies between E5 and E6, again suggestive of regulating alternative splicing.

For SNP rs11576766, AB genotype confers risk over AA and is protective over BB (Table 4.1). The table shows a significant difference between AA and AB. From the figure (Figure 4.4F), AB has a higher methylation than AA, indicating presence of higher methylation state present between E5 and E6, affecting splicing.

Figure 4.4G shows the risk-conferring genotype AB of rs11576766 to have a higher methylation upstream of the gene, suggestive of a down-regulation of *CA6* expression.

SNP rs17032921 shows higher methylation in BB genotype. The risk genotype AA is absent due to the limited number of samples assayed. The graph (Figure 4.4H) can be extrapolated to suggest that the risk genotype (AA) has a much lower level of methylation, as with the presence of the allele A in the heterozygous state is enough to show a statistical difference. Thus, a lower methylation state between E5 and E6 associated with rs17032921 seems to affect splicing in cases.

Our experiments strongly suggest that splicing and gene expression is affected by the associated SNPs. The downstream effect of these processes would affect the translated CA6 protein on the skin.

4.3.3 Protein levels from stratum corneum

Stratum corneum samples were collected from 25 cases and 59 controls. CA6 was measured via immunoblotting. There was no significant difference between cases and controls (Figure 4.5). This could be explained by the different associated SNPs exerting regulatory effects via methylation in opposing trends

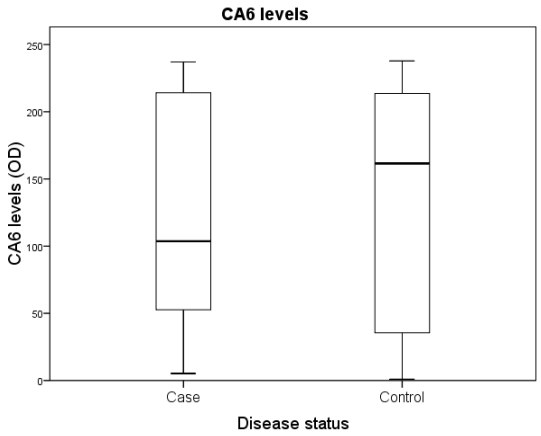

Figure 4-5: CA6 protein levels between cases and controls.

The protein levels were also checked against the genotypes for the associated SNPs. No association was observed. The effect of alleles were also compared to the levels of CA6 and upstream SNP rs1409147 showed mild association (p value 0.046) as shown in Figure 4.6. SNP rs942969 which had a similar effect on the disease status (Table 4.1) showed a similar trend although was not found to be significant.

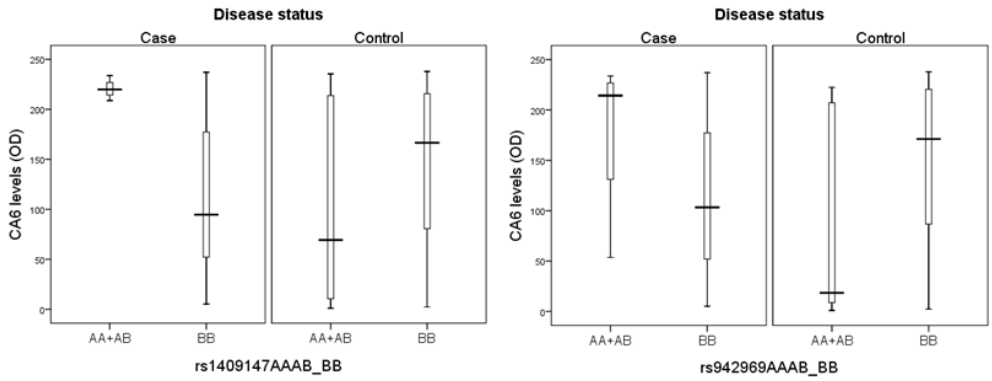

Figure 4-6: CA6 protein levels for rs1409147 and rs942969.

The graph above (Figure 4.6) shows that associated SNPs found upstream (rs1409147) and in intron1 (rs942969) may have an effect on the levels of the protein expressed.

4.3.4 pH and TEWL readings

To evaluate the physical and biological effect of these SNPs, a sub-population was re-called and TEWL and pH readings were measured on 15 cases and 41 controls. The TEWL readings were measured before and after tape stripping. Tape stripping simulates barrier dysfunction and thus we recorded any changes in skin parameters between cases and controls Only the significantly associated SNPs with TEWL and pH readings are described in the Table 4.4.

		rs11576766 AA vs AB	rs7545200 AB vs BB
pH	Pre-tape stripping		0.025
	Post-tape stripping	0.033	
	pH		0.048
TEWL	Post-tape stripping		0.039
	TEWL		0.012

Table 4-4: pH and TEWL associated with significant SNPs.

Table 4.4 shows only the SNPs associated at the different levels of comparisons. pH and TEWL category is the average of the pre and post tape stripping.

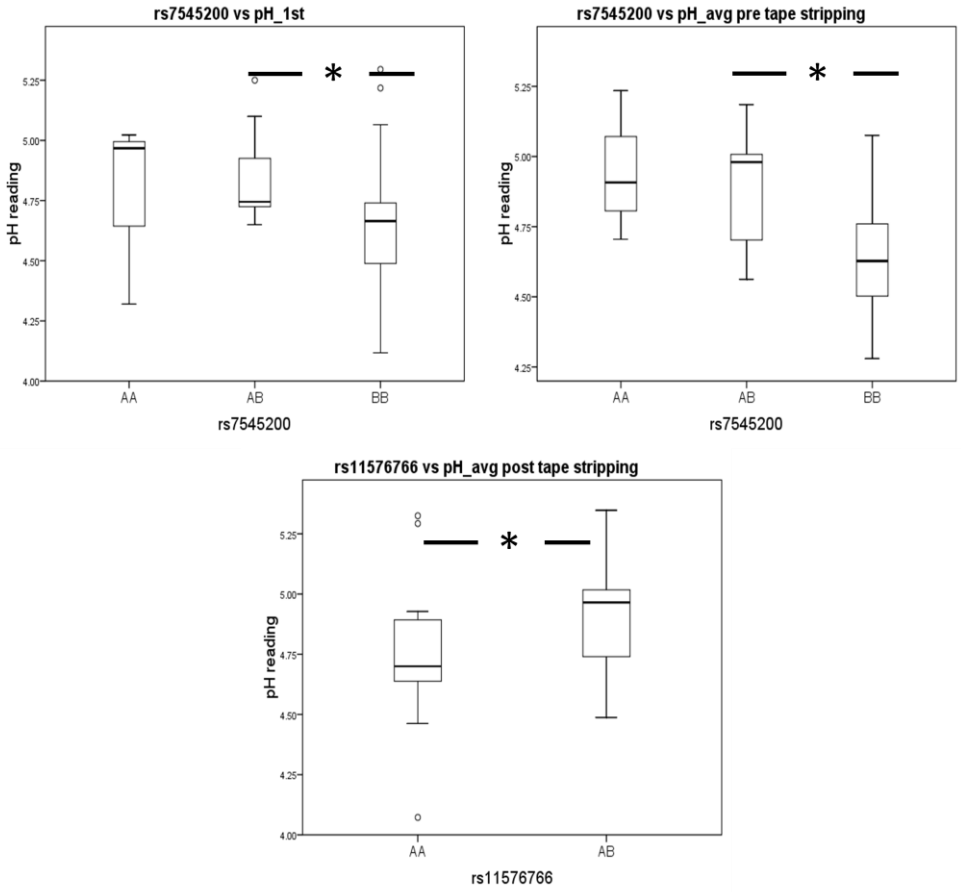

Figure 4-7: pH readings for associated SNPs.

pH readings were found to be different for 2 of the 8 SNPs. rs11576766 showed a difference between AB and AA. AB genotype which is risk-conferring showed an increased pH post-tape stripping. rs7545200 was found to have a higher pH reading in AB as compared to BB before tape stripping and in general. The risk allele A which is present in the heterozygous state showed a consistent high pH reading in both scenarios and in the similar direction as rs11576766.

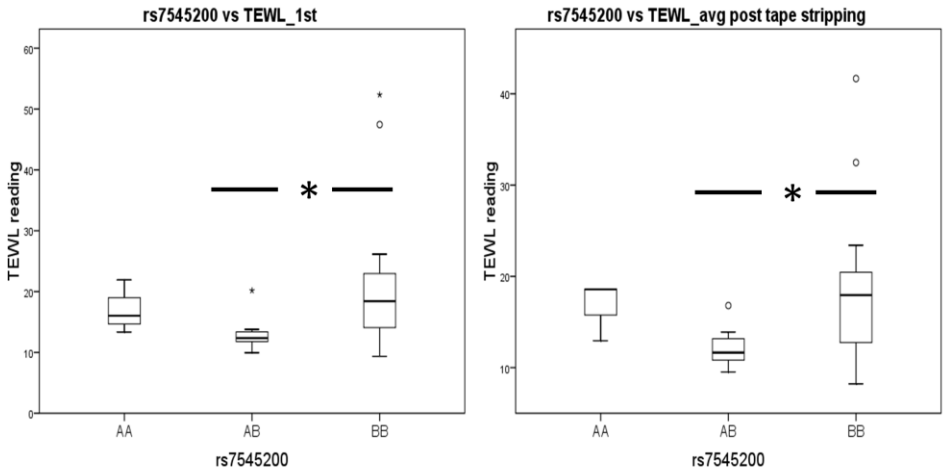

Figure 4-8: TEWL readings for associated SNPs.

TEWL reading was found to be lower in AB as compared to BB. The risk-conferring genotype AB showed lower trans-epidermal water loss in general and post-tape stripping.

4.3.5 *In vitro* experiments

As the significant SNP is a tag, SNPs in LD of over r^2 of over 80% are also considered significant. Table 4.5 gives information on the location of the tag SNP and identities of the SNPs in LD with the significant tag SNPs.

Sig tag SNP	Position	Location	SNPs represented from HapMap
rs1409147	9005320	Upstream	
rs942969	9007556	Intron1	rs12070475,rs6577541,rs3765966
rs2274327	9009406	Exon2	rs12738365
rs2274328	9009444	Exon2	
rs11576766	9010984	Intron2	
rs17032921	9016390	Intron2	rs6577545,rs17389460,rs2274332,rs6691526,rs11805912
rs7545200	9018034	Intron3	
rs3765965	9021998	Intron4	

Table 4-5: Significant tags SNPs from *CA6*.

Of the 8 tags associated, 5 tag SNPs are 'singletons', meaning they represent no other SNP other than themselves. Whereas, the remaining 3 significant

SNPs have SNPs in LD with them which span different regions of the gene as shown in Fig 4.9. As the gene is too large, a portion of the gene crossing the 5th exon is shown in the figure below. To note, E3' is a reported alternatively spliced exon which otherwise would be a part of the intron.

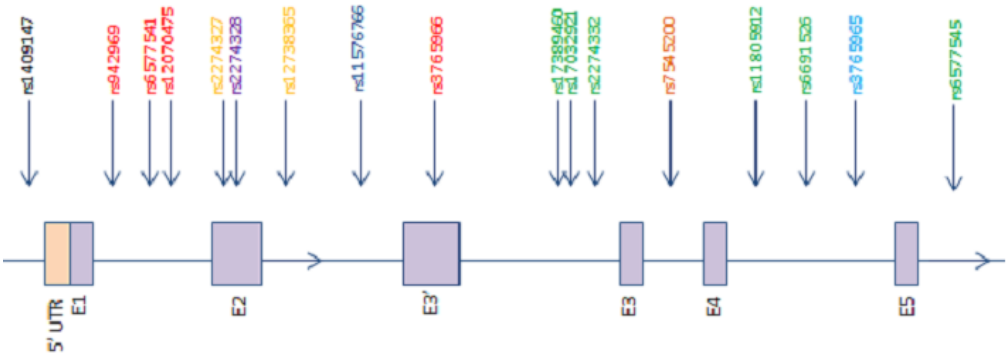

Figure 4-9: Significant tags and associated SNPs of *CA6*.
Abbreviations: 5' UTR- Untranslated region, E- Exon.

4.3.5.1 Functional predictions

Functional predictions for the significantly associated tags and SNPs in LD was conducted using the databases as mentioned in section 4.2.6. Depending on the location, SNPs were probed for different functions. The compiled function based on the various searches is tabulated below (Table 4.6).

			SNPInfo			F SNP		Transcription factor binding		Splicing
SNP ID	Location	Allele	TFBS	Splicing(ESE or ESS)	nsSNP	Region	Score	TFSearch	ConSite	HSF
rs1409147	Upstream	A/G	Y	-	-	Intronic	-	-	-	
rs942969	Int1	C/G	-	-	-	Intronic	-	-	SOX 17, AML-1	Y
rs6577541	Int1	C/G	-	-	-	Intronic	-	GATA-1: 91.6, MZF1: 86.1	-	Y
rs12070475	Int1	C/G	-	-	-	Intronic	-	CdxA 92.9, TATA 87.8	-	Y
rs2274327	E2	C/T	-	-	Y	Non-syn (T55M), Splicing, PTM	0.909			Y
rs2274328	E2	A/C	-	Y	Y	Non-syn (M68L), Splicing	0.33			Y
rs12738365	Int2	C/T	-	-	-	Intronic (Tx regulation)	0.101			Y
rs11576766	Int2	A/C	-	-	-	Intronic	-			Y
rs3765966	E3'	T/G	-	Y	-	Non-syn (R-->L), Splicing	0.917			Y
rs17389460	Int 3'	G/T	-	-	-	Intronic	-			Y
rs17032921	Int 3'	C/T	-	-	-	Intronic (Tx regulation)	0.101			Y
rs2274332	Int 3'	C/T	-	-	-	Intronic (Tx regulation)	0.101			Y
rs7545200	Int3	A/G	-	-	-	Intronic	-			Y
rs11805912	Int4	C/T	-	-	-	Intronic (Tx regulation)	0.101			Y
rs6691526	Int4	A/G	-	-	-	Intronic (Tx regulation)	0.101			Y
rs3765965	Int4	T/C	-	-	-	Intronic	-			Y
rs6577545	Int5	C/G	-	-	-	Intronic	-			Y

Table 4-6: Functional predictions based on location of the SNPs.
Abbreviations: E- Exon, Int- Intron, Y- Yes/ Present.

Based on the above table, two regions i.e. the upstream and first intron were selected to probe for checking promoter-like activity. Due to the size limitation of the insert in the plasmid, two regions were selected for affecting splicing. The first region spanned exons 1,2,3',3,4 and the second construct spanned exons 2,3',3,4,5. As splicing was being checked by the construction of a minigene, exons along with 200 - 300 bases of the introns flanking the exons were joined together.

4.3.5.2 Resequencing

In HapMap, no SNPs with a MAF of 0.05 and above were found around the significant upstream tag (rs1409147) and in the introns flanking the second exon. Thus, resequencing was undertaken for this region to identify the different haplotypes present in our population.

A 5 kb region was successfully amplified using the long range high fidelity Kapa HiFi polymerase and areas of interest were sequenced. The sequencing yielded the following LD pattern spanning upstream of the gene and from the first intron into the second intron (Figure 4.10). SNPs number 1-4 were upstream of the gene and SNP numbers 5-11 spanned intron 1, exon 2 and intron 2. SNP number 4 labelled as 'SNP1' was a novel SNP found, whereas the remaining 10 SNPs were reported SNPs in NCBI's SNP database.

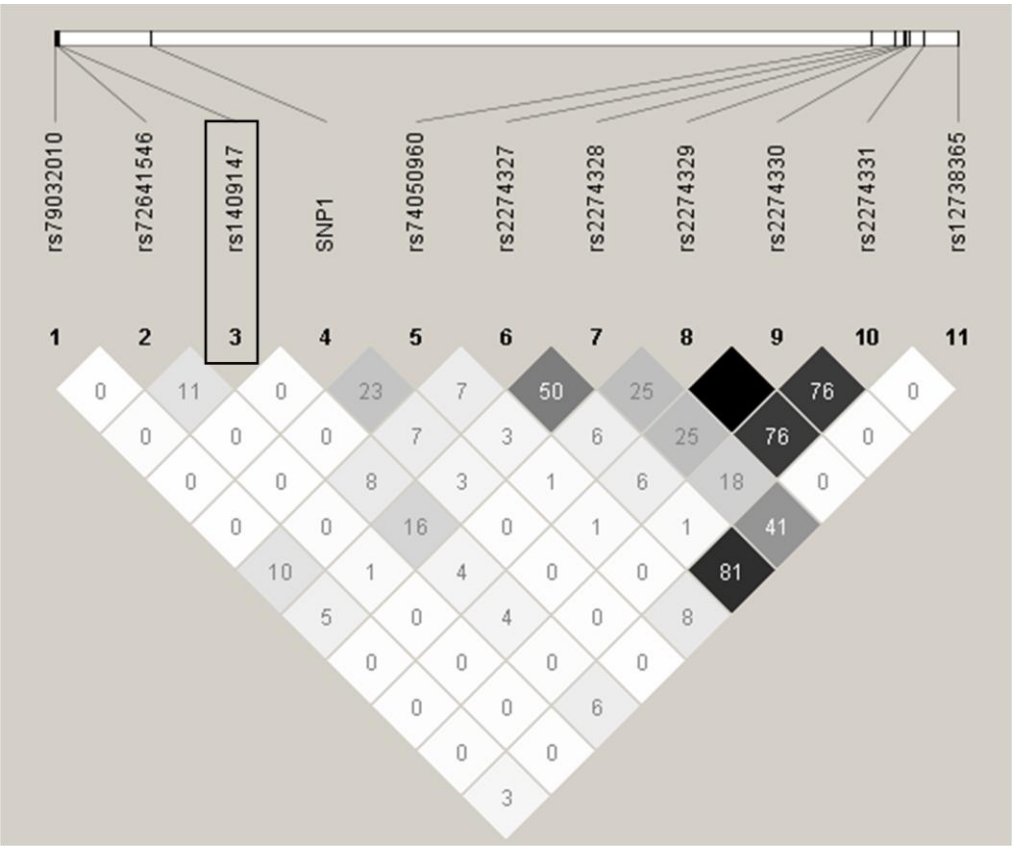

Figure 4-10: LD pattern found through re-sequencing.

The numbers in the intersection of 2 SNPs represents the LD value in r^2. Absolute black squares indicate perfect LD, i.e. r^2 value of 100% between the 2 SNPs as found in the case of SNP 8 (rs2274329) and SNP 9 (rs2274330).

4.3.5.3 LD and haplotypes

The upstream haplotype around SNP rs1409147 (SNP 3) identified through re-sequencing is as shown in Fig 4.11

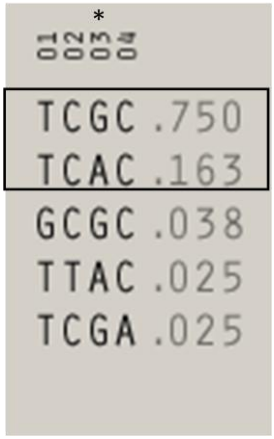

Figure 4-11: Upstream haplotypes.

The numbers 1 to 4 match with the previous figure (Figure 4.10). The boxed-haplotypes were selected for *in vitro* assay as their frequencies were above 5%.

The haplotypes for checking intron 1's promoter activity are as shown in Figure 4.12. SNP rs10864372 (SNP 5) was found to be present in SGVP in 100% LD with the significant tag rs942969 and was thus incorporated in the construct. rs74050960 (SNP 6) found through sequencing was also included.

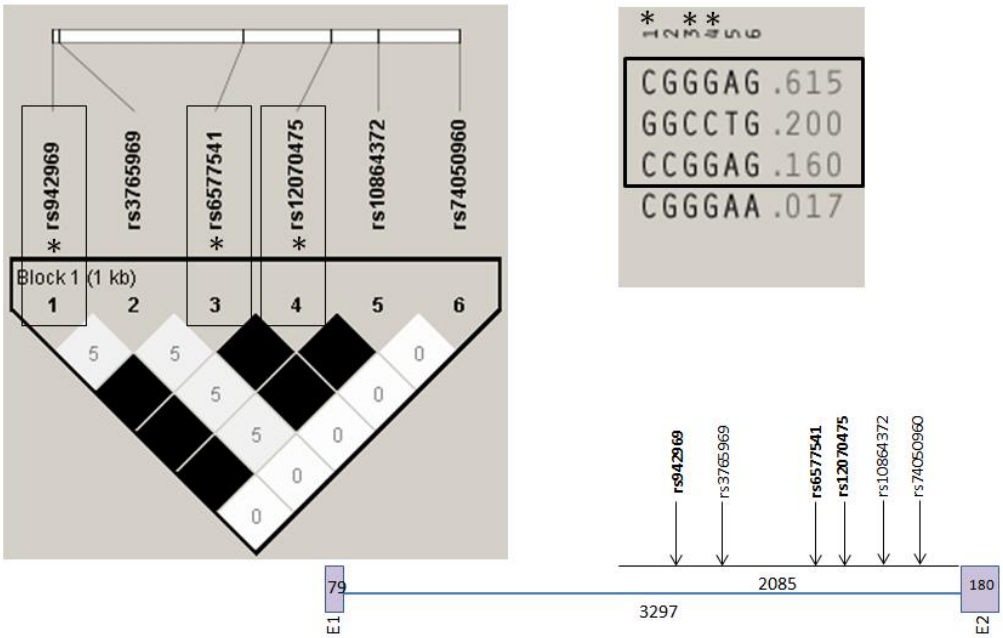

Figure 4-12: LD pattern and haplotypes for the Intron 1 construct.

The intron 1 construct for checking promoter activity was made with 6 SNPs as shown in Figure 4.12. As highlighted, three haplotypes present at a frequency of more than 5% were selected for *in vitro* activity.

4.3.5.4 Luciferase reporter assay for promoter

Two haplotypes for upstream construct probing for rs1409147 and three haplotypes for intron 1 construct probing for rs942969 were put through luciferase reporter promoter assay. The ratio of luminescence (normalised) was plotted. Vc represent the vehicle control which is the empty pGL4.10.

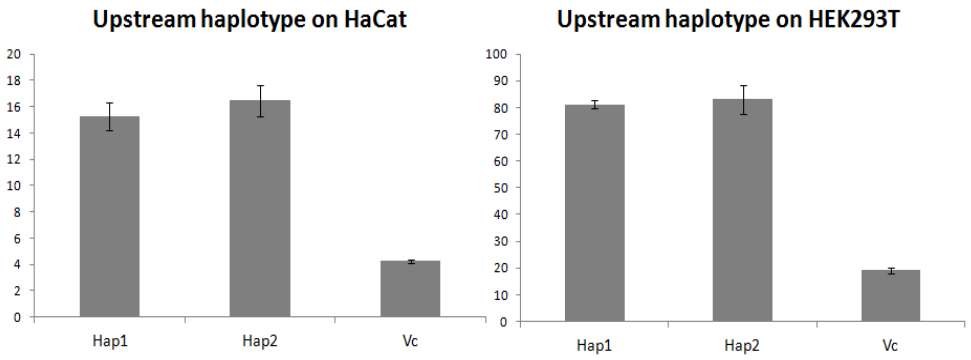

Figure 4-13: Luciferase promoter assay for upstream haplotypes.

The upstream construct showed promoter activity when compared to the empty vector in both cell lines. However, there was no difference observed between the two haplotypes (Figure 4.13).

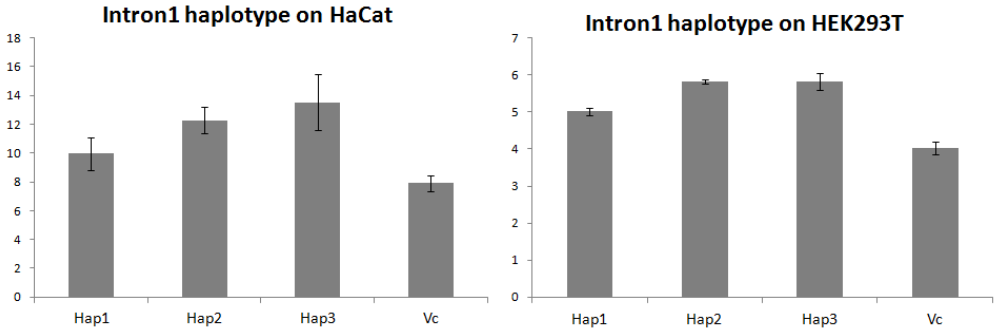

Figure 4-14: Luciferase promoter assay for intron 1 haplotypes.

The intron 1 construct did not show any promoter activity as the three haplotypes have intensities similar to the vehicle control. Also, there was no difference between the three haplotypes (Figure 4.14)

Thus the two regions probed for promoter activity showed no difference between haplotypes in the two tested cell lines.

Similarly, to check for splicing around the other associated SNPs, 2 constructs were designed and ordered.

4.3.5.5 Minigene experiment for alternative splicing

For the splicing minigene experiment, construct 1 spanning E1- E4 was designed as shown below. Eight haplotypes for this construct (minigene 1) were ordered as they were above 5% MAF.

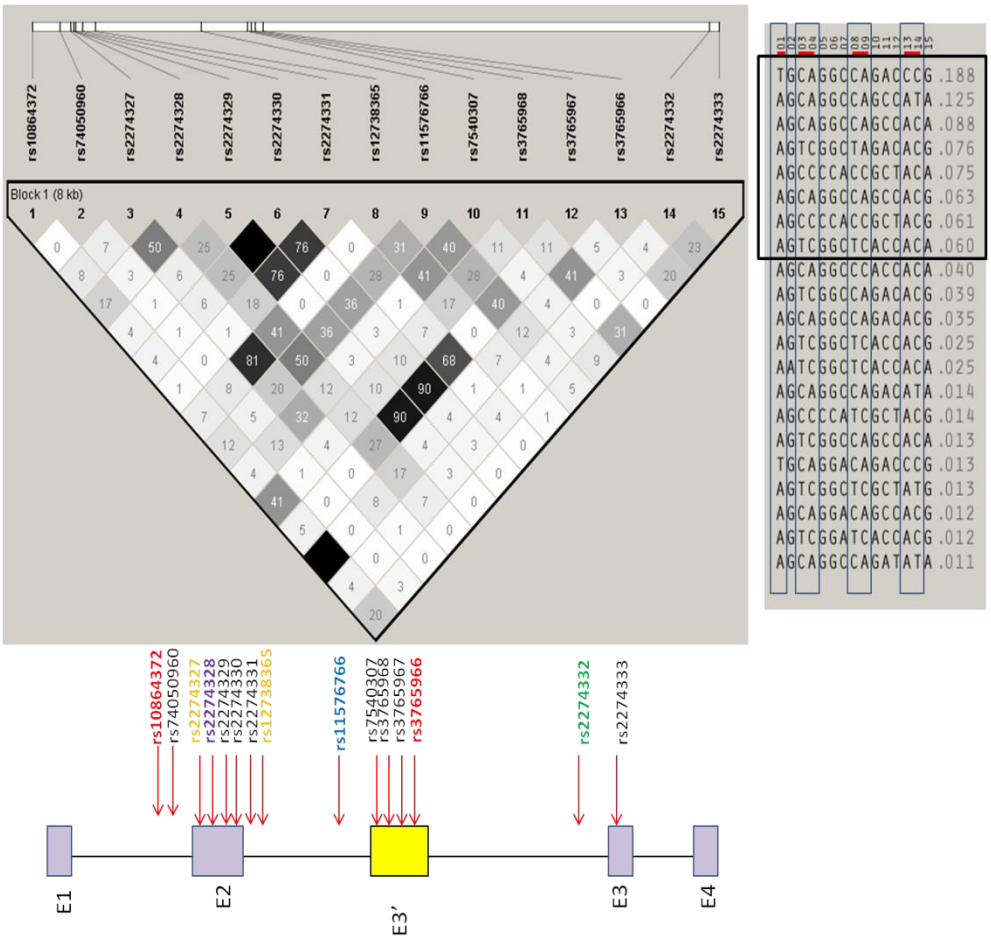

Figure 4-15: LD, haplotypes and construct for minigene1.

Similarly, another construct spanning E2- E5 was designed as shown below in figure 4.16. Eight haplotypes for this construct (minigene 2) were ordered as they were above 5% MAF.

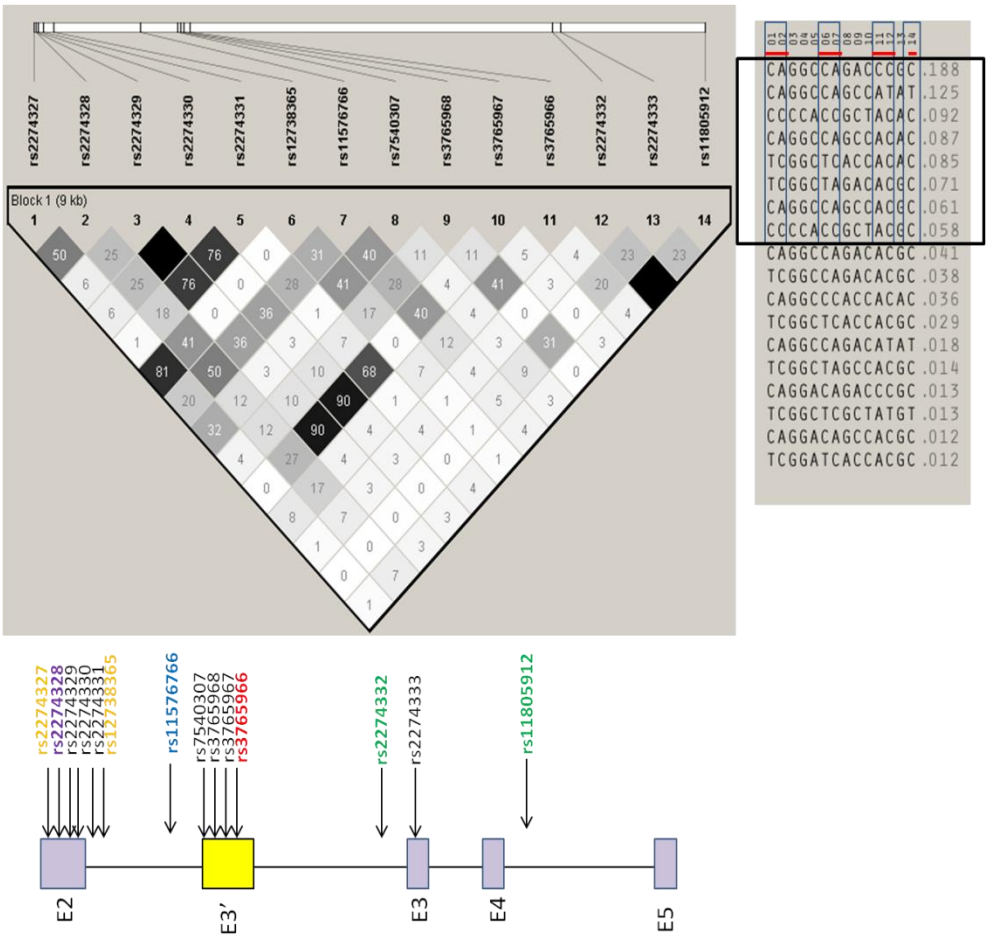

Figure 4-16: LD, haplotypes and construct for minigene2.

In vitro mini-gene experiment for splicing were conducted in HaCat and HEK2983T for both constructs with 8 haplotypes each. Experiment was done in triplicates, however, data was inconclusive (data not shown).

4.4 Discussion

CA6 had 8 tags associated with the disease. SNPs in LD with the significant SNPs ranged from upstream till after exon 5. To characterise the effect of these tags, exon array, methylation chips were run. Protein analysis and

physiological parameters of the skin were also investigated with respect to these SNPs.

4.4.1 Alternative splicing

More than 85% of SNPs found associated through GWAS and candidate gene studies are located in the introns or intergenic areas. SNPs in exonic areas are extremely few and with even fewer of those associated lead to change in amino acid. Even with a change in amino acid, extremely few of them actually lead to a drastically different amino acid with the possible outcome of altered protein function and stability. Similar trend of an overpowering number of associations found in intronic regions over exonic regions in our study corroborates the observed trends in literature. As a result, the large number and strong associations of intronic regions suggests a causal role in gene regulation to disease pathogenesis (Costa *et al.*, 2013). A sizeable number of SNPs associated with the intron-exon boundary and exons (synonymous and non-synonymous) affect splicing, causing diseases (Krawczak *et al.*, 2007). Given the diversity of known and still to be known mechanisms involved in splicing, alternative splicing may play a bigger role in disease pathogenesis. Results for gene expression and selective exon over-expression have been discussed in the context of the associated SNP in section 4.4.6.

4.4.2 Regulation by DNA methylation

DNA methylation is one of the major epigenetic regulator of gene expression. DNA methylation occurs in around 60-90% of the Cytosine-Guanine dinucleotides in mammals (Siegfried and Cedar, 1997). The link of DNA methylation and gene silencing is largely published and well accepted (Kass *et*

al., 1997; Keshet *et al.*, 1985; Yisraeli *et al.*, 1988). This repression is suggested to be due to the interference in the recognition and binding of transcription factors to the methylated DNA (Eden and Cedar, 1994). However, there are transcription factors which bind independent of the methylated state of the recognition sites. Another model suggests that methylated DNA attracts proteins which blocks access of different transcription induction factors (Boyes and Bird, 1991). Overall, a physical block placed either by the methyl-group or a protein bound to the methylated region causes repression.

More recently, methylated regions found within the gene have been implicated in alternative splicing. Alternative splicing occurs when the weak splice sites compete with the stronger splice sites for being detected by the splicosome machinery (Matlin *et al.*, 2005). This competition is regulated by (i) the presence of splicing factors which can detect enhancing and silencing sequences on the RNA and (ii) the rate of RNA polymerase II elongation (Han *et al.*, 2011). Epigenetic markers such as methylation add complexity to the DNA structure affecting the co-transcriptional processing of RNA. Recent studies have shown contrary effects of methylation such as (i) preventing binding of splicosome factors thereby causing exon skipping (Shukla *et al.*, 2011) and also (ii) enhancing exon recognition resulting in exon inclusion (Maunakea *et al.*, 2013).

With this in view, it can be concluded that methylation found upstream leads to transcription repression whereas methylation found within the gene could either promote in including or excluding the near-by exons. Results for

methylation associated gene repression and splicing have been discussed in the context of the associated SNP in section 4.4.6.

4.4.3 CA6 protein levels

Protein level difference were evaluated between 25 cases and 59 controls. No significant difference was observed between the two groups (*P* value = 0.780), however, cases did show a trend of lower levels in cases as compared to controls when observing the median levels (Fig 4.5). Protein levels were analysed with reference to genotypes of the associated SNPs. SNP rs1409147 associated at the allelic level and lying upstream of the gene showed marginal association to protein levels (*P* value = 0.046) within cases. To further visualise the difference, data was plotted in a case-control manner (Fig 4.6). In the figure, the risk conferring allele B showed lower levels of protein in cases, while the same group had higher levels in controls. As AD is a complex and non-Mendelian disease, the effect of risk conferring allele is not absolute as depicted in Fig 4.6. Samples containing the risk conferring allele B with higher levels of the proteins constituted as controls, probably due to the protective alleles/genotypes of other associated SNPs. A similar trend was observed for SNP rs942969, however the difference was not significant (*P* value = 0.177). Protein levels in context with the associated SNPs and methylation has been discussed with greater detail in section 4.4.6.

4.4.4 pH effect

The acid mantle of the skin is a key regulator of maintaining homeostasis (Marchionini and Hausknecht, 1938). The usual pH is in the acidic region of 5.4 - 5.9, with differences observed between gender, ethnicity and region of the skin (Braun-Falco and Korting, 1986). Increase in pH has been shown to

be associated with eczema (Eberlein-Konig *et al.*, 2000) and also cause barrier dysfunction (Hachem *et al.*, 2003). Serine proteases involved in desquamation have an optimum activity at neutral pH (Ekholm *et al.*, 2000), thus the acidic pH keeps the activity under control. An increased pH could send these proteases in hyper drive causing barrier dysfunction. Increased pH also reduces the activity of lipid synthesizing enzymes (Holleran *et al.*, 1993). Also, methods such as tape stripping have been known to disrupt the barrier (Bashir *et al.*, 2001) causing an increase in pH (Dikstein and Zlotogorski, 1994; Ohman and Vahlquist, 1994). Moreover, increased pH is shown to delay barrier recovery (Mauro *et al.*, 1998). Results for the pH association data to the significant SNPs has been discussed in section 4.4.6.

4.4.5 TEWL

Trans-epidermal water loss is a measure of the permeability barrier of the skin. Higher TEWL reading indicates increased rate of evaporation (Fluhr *et al.*, 2006). This can be due to an inherent defective permeability barrier or due to external physical stress such as tape stripping or scratching. Increased TEWL readings has been shown to be negatively correlated with skin hydration and positively correlated to a few proteases of the SC (Voegeli *et al.*, 2008). TEWL is also shown to be associated with thickness of the barrier (Nikolovski *et al.*, 2008). Results for TEWL association data to the significant SNPs has been discussed in section 4.4.6.

4.4.6 SNPs

With the above information and discussion in view, the associated SNPs are discussed below.

4.4.6.1 rs1409147

The upstream SNP rs1409147 is associated with the disease state through allele B. Exon array data showed the BB genotype had reduced expression of the probe in the upstream region. This could imply a reduction in the transcription in presence of BB. Also, CA6 protein levels showed that BB in cases had significantly low levels as compared to AA+AB genotypes, and this trend is reversed in the control group. These results hint towards a reduced expression of *CA6* when associated with BB genotype of rs1409147. When analysing methylation differences between case and control, methylation site upstream of the gene was found to be slightly more methylated in cases than in controls. Methylation upstream is correlated with gene silencing or reduced gene expression. Functional prediction of the SNP had shown presence of transcription factor binding site thereby strengthening the association of rs1409147 to the expression levels of *CA6*. *In vitro* luciferase reporter experiment also showed the presence of a promoter activity, although no difference between the haplotypes was observed. Even though *in vitro* experiments did not show a difference between the haplotypes, the remaining data strongly suggests the risk allele of rs1409147 to have reduced overall expression of *CA6*. *In vitro* demonstration may have failed due to the possible lack of the specific transcription factor which might have the ability to differentiate the alleles.

4.4.6.2 rs942969

rs942969 present in intron 1 is associated with AD by B allele and BB genotype as compared to AA. rs942969 is in LD with SNPs further in the intron 1 and 1 SNP in intron 2 which maps onto an alternatively spliced exon referred to as E3' in this chapter. Functional prediction for the rs942969 and the SNPs in LD with it showed putative regulation by transcription factor and a possible splicing effect. To evaluate the functional prediction, haplotypes present in the first intron were tested through luciferase promoter assay. However, the construct did not shown any promoter activity as compared to the vehicle control. With this, it can be concluded that rs942969 is not associated through transcription based regulation. Methylation based study revealed low methylationin in the region E1 associated with the risk-conferring BB genotype. Methylation with the gene has been shown to affect splicing. Thus, it can be concluded that rs942969 regulates *CA6* via methylation. However, the exact effect of the methylation still needs to be ascertained.

4.4.6.3 rs2274327

SNP rs2274327 present in exon 2 is associated with AD through the A allele and through AB genotype when compared against BB. Besides a predictive effect on splicing, rs2274327 causes a change of amino acid from Threonine to Methionine at 55th amino acid position. On discussing with a structural biologist, it was suggested that the variant lying on an exposed helix may alter the stability of the protein. Protein stability of the two variants in different salt and pH conditions are planned to be checked through circular dichroism in the near future. Methylation study found the probe present between E6 and E7 to

be more methylated with the risk genotype of AB indicating a possible role in splicing.

4.4.6.4 rs2274328

rs2274328 is associated with AD through allele B and through AB genotype when compared with AA. This SNP also lies in exon 2 and causes a change from Methionine to Leucine at amino acid position 68. Structural evaluation would be needed to check if the change affects the structure or activity in anyway. The risk-conferring genotype AB was shown to be associated with *CA6* over-expression. AB was also found to be associated with increased transcription of exon 1 and exon 9. Methylation study also showed AB was associated with higher methylation in the region between E5 and E6 indicating its possible role in splicing.

4.4.6.5 rs11576766

SNP rs11576766 shows a similar trend as rs2274328. The AB genotype has a risk effect as compared to AA and BB and was shown to have higher gene expression in the risk genotype group. AB was also found associated with enhance transcription of upstream, exon 9 and 3' UTR. The AB genotype showed significantly higher levels of methylation as compared to AA and a trend of higher methylation as compared to BB for the upstream methylation site. This suggests, that reduced gene expression as a result of methylation of upstream regions is associated with AD cases. This follows the same trend of the upstream SNP rs1409147. Higher methylation between E5 and E6 is observed in AB as compared to AA again suggesting a possible role in splicing of the exons around that region. Thus rs11576766 seems to regulate gene in two very different ways via methylation. Another interesting

observation was the increased pH after a physical stress in the AB genotype. This suggests an inherent aberration in the processes involved in maintaining homeostasis as in AD. Based on literature explained earlier, this group would have a slower recovery leaving their skin open to allergens and pathogens for a longer time thus predisposing to AD.

4.4.6.6 rs17032921

rs17032921 lying in the intron 2 of the gene represents 5 SNPs based on LD. The SNPs in LD are present in intron 2, intron 4 and intron 5. rs17032921 is associated with AD through allele A and genotype AA as compared to BB. Although predicted to have an effect on splicing, there was no differential expression of exons observed. In the methylation study, no sample with AA genotype was assayed due to the small numbers. However, AB showed significant lower methylation as compared to BB for the regions between E5 and E6 (intron 5). This indicates a possible regulation of alternative splicing due to methylation.

4.4.6.7 rs754200

SNP rs7545200 present in intron 3 is a singleton SNP associated with AD through the allele A. This SNP was found to be associated with increased pH in general and with increased pH prior to tape stripping in the AB genotype over BB, and AA+AB versus BB genotypes (data not shown). The increased pH as with rs11576766 is a marker for dysfunctional barrier. The risk group (AA+AB) has a higher pH as compared to BB similar to that observed in people with AD. The risk conferring group (AA+AB) showed a marginal decrease in TEWL as compared to BB genotype. Although increased TEWL is

the hallmark of AD, it is highly variable upon ambient conditions and physical condition of the skin due to scratching or rubbing.

4.4.6.8 rs3765965

rs3765965 lying in the intron 4 is associated with AD through BB genotype as compared to AA. No samples from the minor genotype BB were put through methylation, thus an analysis of AB versus AA was done. AB genotype showed lower methylation associated with the area between E5 and E6 as compared to AA. This could link the putative function of rs3765965 splicing regulated by methylation.

Overall, all the associated 8 tags were attributed to some function thereby strengthening the results of the association study. Different SNPs contributed to the pathogenic model in different ways. For eg: rs1409147, differences between case-control and upstream methylation by rs11576766 showed suppressed expression of *CA6* to be associated with AD. However, gene expression associated with rs2274328 and rs11576766 show increased expression in the risk genotype. This shows, different mechanisms at play having contradictory effect. This may be the reason why no difference in CA6 protein levels were observed between the well powered case-control study.

4.4.7 Carbonic anhydrase VI

CA6 also known as gustin is a secreted form of carbonic anhydrase. It is a zinc metalloprotein which is thought to catalyse the hydration of carbon dioxide in regulating pH (Carter, 1972; Khalifah, 1971b).

$CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$

Carbonic anhydrases are involved in maintaining cellular pH, water transport and ion homeostasis (Nakhoul *et al.*, 1998). CA6 is abundantly found in saliva (Fernley *et al.*, 1979), milk (Karhumaa *et al.*, 2001), gut (Parkkila *et al.*, 1994) and other epidermal tissues. It is hypothesized to have a mucosa protective role in the gastro-intestinal tract of neonates, as the colostrum has 8 times more CA6 than mature milk and saliva (Karhumaa *et al.*, 2001). It is also thought to be a growth factor for taste-buds. SNP rs2274327, significant in our study, has been shown to increase buffering capacity of the saliva (Peres *et al.*, 2010).

CA2 a member from the same family was shown to be up-regulated in AD skin biopsies when compared with psoriatic and normal skin. The group also found CA2 to be inducible by Th2 cytokines in keratinocytes (Kamsteeg *et al.*, 2007). The up-regulation of CA2 doesn't seem to be genetic as we found no association for *CA2* in our genetic study. Also, inhibition of CAs has been shown to lower intracellular pH (Kniep *et al.*, 2006).

A fine balance in the pH is required to optimally activate the proteases and their inhibitors, an aberration in which can over process the inactive proteases thereby causing barrier dysfunction. Proteases, besides degrading the corneodesmosomes and lipid producing enzymes, also aid in the downstream over-expression of pro-inflammatory cytokines such as IL-1α, IL-1β, TNF- α and IL-6 (Wood *et al.*, 1996; Wood *et al.*, 1997). pH is also believed to play an important role in maintaining the normal flora of the skin by favouring adhesion of non-pathogenic bacteria (Bibel *et al.*, 1987) and preventing colonisation of pathogenic microbes (Rebell *et al.*, 1950). Thus, pH seems to

be the master regulator of all the key processes involved in maintaining the epidermal barrier. Consequently, a novel association of a possible pH regulating gene (*CA6*) to the pathogenesis of AD is an important discovery and paves the way for more detailed therapeutic research of pH control.

Since its discovery in 1940s, not much is known about the true function of CA6 (Sly and Hu, 1995). Its presence, regulation and similarities with other members have resulted in most of the hypotheses for its function. This study in addition with the other cited above further stress the importance for detailed research of this gene.

4.5 Conclusion

Association of the *CA6* to the disease was established through genotyping study. Association of the significant SNPs to methylation- regulated and alternative splicing was also established. However, *in vitro* characterisation of the associations still needs more work. Defined gene function of *CA6* which is still elusive after years of discovery would be needed in order to bridge the effect of the SNPs back to disease pathogenesis.

Chapter 5

Replication studies

5.1 Validation of Han Chinese AD GWAS in Singapore population.

Part of this chapter is published in *Journal of Investigative Dermatology* (2012) titled, 'Validation of GWAS loci for atopic dermatitis in a Singapore Chinese population'

5.1.1 Introduction

Candidate gene study approach for identifying genes associated with a disease relies on *a priori* hypothesis for the disease model. As in this thesis, pathways or members of protein families or genes expressed and localised at the disease site are few of the starting points for initiating a candidate gene study. However, the genome wide association studies (GWAS), is a hypothesis generating study wherein no prior information on likely candidates is needed. Study design and a well powered sample size are the initial stages of a GWAS with the results giving leads to the possible pathways at play. With high throughput and low cost, GWAS for complex diseases is affordable to labs.

One of the major bottle-necks to a successful GWAS are large numbers. Power curves have shown that to obtain a modest effect size of OR 1.3 for SNPs with a MAF of less than 0.1, sample sizes in thousands is needed (Wang *et al.*, 2005). The large number of probed SNPs results in an even higher rate

of false positives. To be able to discern the difference between the true and false positives (Wacholder *et al.*, 2004), multiple corrections for such a high number of probed SNPs leads to a p value cut off that many studies fail to cross.

GWAS for AD have yielded novel candidates in the recent past. In 2009, a new candidate on chromosome 11q13.5, *C11orf30* was discovered in European populations. The study's discovery set comprised of 939 cases and 975 controls and 270 nuclear families and the significant SNP was validated on 2637 cases and 3957 controls. The study found a SNP rs7927894 which was 38kb downstream to *C11orf30.* Around 13% with an European ancestry carry the homozygous variant of this SNP and showed an OR of 1.47 (Esparza-Gordillo *et al.*, 2009). GWAS in Japanese population identified 8 new loci - *IL1RL1-IL18R1-IL18RAP,* MHC region, *OR10A3-NLRP10, GLB1, CCDC80, CARD11, ZNF365* and *CYP24A1-PFDN4* from a population of 3328 cases and 14992 controls (Hirota *et al.*, 2012). Another discovery study on 1536 cases and 4054 controls and a validation on 2286 cases and 3160 controls of European ancestry resulted in regions; (i) of the epidermal differentiation complex on chromosome1, (ii) around LRRC32 on chromosome 11, (iii) around *RAD50/IL13* locus on chromosome 5 and (iv) around MHC on chromosome 6 (Weidinger *et al.*, 2013). GWAS conducted on canines have also resulted in interesting candidates such as *SORCS2, RAB3C* and *PROM1* (Wood *et al.*, 2009) and *PKP2* (Tengvall et al., 2013).

The only GWAS on a Chinese (Han Chinese) population was published in 2011. The initial study was on 1012 cases and 1362 controls which was then validated on a 3624 cases and 12197 controls of the same ethnicity and

replicated on 1806 cases and 3256 controls from Germany. They found and validated 2 new loci in chromosome 5q22.1 (rs7701890) and 20q13.33 (rs6010620). They replicated the 1q21.3 locus (*FLG* rs3126085) and provided suggestive association at 10q21.2 (rs2393903) (Sun *et al.*, 2011). With the inability to match their sample size and a close genetic link of the Singapore Chinese with ethnic Han Chinese, we decided to validate the 3 SNPs from the Chinese AD GWAS results in our population.

5.1.2 Materials and method:

The samples used for this study overlapped with those described in Chapter 2 and Chapter 3 (NUS + KKH + IMB/NSC). In addition 400 Singapore Chinese controls were ascertained from Bio Bank as population controls (Common *et al.*, 2011). A total of 827 AD cases and 1104 controls were genotyped.

Genotyping for the 3 SNPs was conducted by using TaqMan assays according to manufacturer's protocol.

5.1.3 Results:

The genotyping results are summarised below. The chromosome 10 (rs2393903) and chromosome 20 (rs60110620) SNPs were found associated with AD in our population. The effect sizes as described by the OR are similar to the Chinese GWAS study. SNP found on chromosome 5 (rs77018980) however, was not found to be associated in our population.

			All AD		AD only		AD + others	
			827 Cases, 1,104 controls		251 Cases, 1,104 controls		572 Cases, 1,104 controls	
Chr	SNP	MAF controls	OR (95% CI)	*P* value	OR (95% CI)	*P* value	OR (95% CI)	*P* value
5	rs7701890	0.20 (G)	0.956 (0.809–1.129)	0.594	1.107 (0.870–1.416)	0.411	0.887 (0.732–1.073)	0.218
10	rs2393903	0.42 (C)	1.264 (1.109–1.440)	0.0004	1.447 (1.210–1.791)	0.0002	1.19 (1.028–1.378)	0.021
20	rs6010620	0.28 (G)	1.179 (1.022–1.360)	0.024	1.045 (0.842–1.303)	0.691	1.244 (1.061–1.459)	0.007

Table 5-1: Case control association for AD GWAS SNPs.

Data was also analysed by splitting the AD cases into those with AD alone (AD only) and those with either asthma or AR or both (AD+ others), as defined in Chapter 2. Results showed a significantly stronger association of risk C allele of the chromosome 10's SNP in AD alone when compared with AD and other atopic conditions (OR 1.45 vs 1.2).

On a subset of the population collected from National Skin Centre through IMB, severity scores were collected through SCORing Atopic Dermatitis (SCORAD) (Oranje *et al.*, 2007).

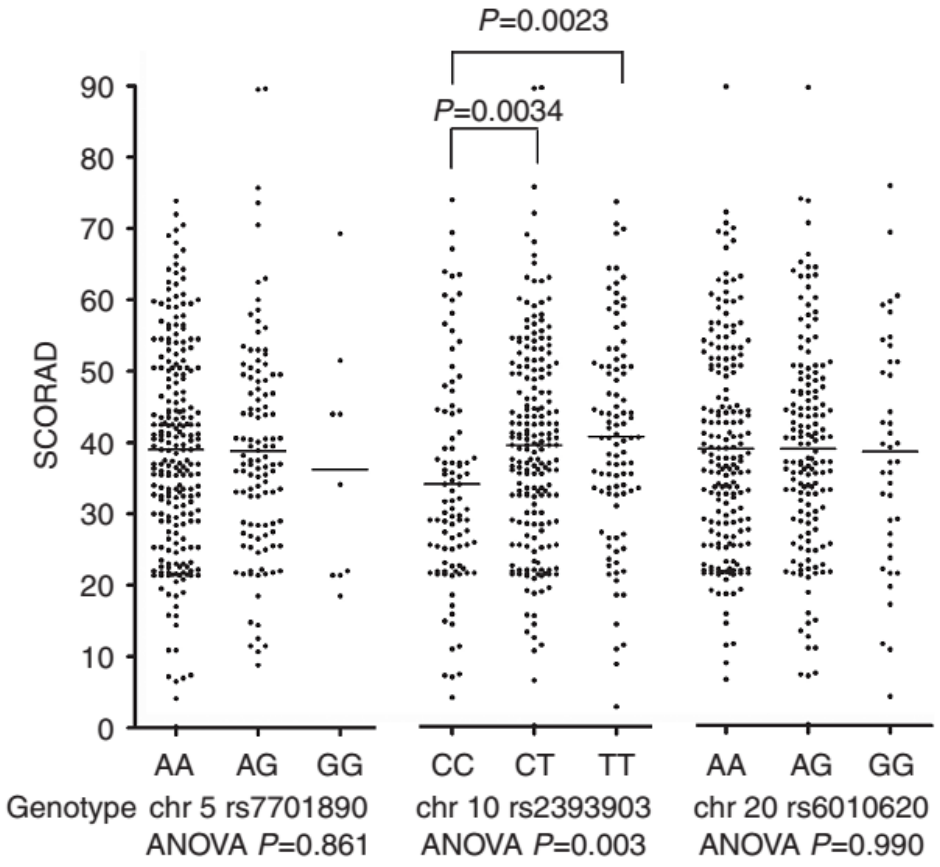

Figure 5-1: Distribution of SCORAD scores vs genotypes.

ANOVA analysis showed a difference for chromosome 10 SNP (rs2393903) with the CC genotype group (mean 33.99±1.568; n=94) having lower SCORAD scores as compared to CT (mean 39.48±1.043, n=197; t-test, *P*=.0034) and TT (mean 40.66±1.491, n=98; t-test, *P*=0.0023) (Figure 5.1). The results suggest rs2393903 is associated with milder form of AD whereas the other SNPs showed no association to severity based on SCORAD scores.

In order to evaluate the lack of association on chromosome 5 (rs7701890), we tested the various parameters evaluated in Chapter 2 using logistic regression on the samples collected from NUS (432 cases and 700 controls). We observed a significant association with household income (*P*= 0.017), in which the minor allele G showed the trend of a risk effect [OR = 1.22 (95% CI

0.92–1.61), *P*=0.17] in the low-income group (<S$4,000 a month) and a protective trend [OR = 0.71 (95% CI 0.49–1.01), *P*= 0.06] in high-income group (> S$4,000 a month).

5.1.4 Discussion

We validated the chromosome 10 (rs2393903) and chromosome 20 (rs60110620) SNPs' association in our population, however, rs7701890 from chromosome 5 (rs7701890) could not be replicated. Previous studies in the Asian region have shown a reduced prevalence of allergic diseases in lower income groups in Singapore (Goh *et al.*, 1996) and Hong Kong (Lau *et al.*, 1995). However, more recently, increased prevalence of eczema in lower income had been reported in the United States (Hanifin and Reed, 2007). This highlights the change and impact of environmental factors that can become possible confounders in future studies. We speculate that higher percentage of subjects from a high-income group in the Singapore Chinese population compared with the subjects from the Chinese GWAS study may explain the lack of association on chromosome 5.

5.1.5 Conclusion

We validated the reported strong association at chr10q21.2 and a weaker one at 20q13.33 with AD among Singapore Chinese, but suggest the genetic association at 5q22.1 might be influenced by the interaction with environmental factors. The validation of new AD loci will help to further understand the genetic basis of this complex disease. However, a much larger sample size would need to be examined to scrutinise the effect of non-genetic interaction to disease (Andiappan *et al.*, 2012).

5.2 Association of *FLG* mutations

5.2.1 Introduction:

As indicated in section 1.7, *FLG* is the most widely associated gene to atopic dermatitis (Palmer *et al.*, 2006). It forms a part of the tough barrier by collapsing the aggregated fibres forming the flattened large surface of the corneocytes (Candi *et al.*, 2005). The FLG protein also deaminated to form free amino acids which are then catabolized into the constituents of water-retaining natural moisturising factor (NMF). Water retention within corneocytes results in their optimal hydration and swelling, which prevents the development of gaps thereby enhancing the integrity of the SC and making it resistant to the penetration of irritants and allergens (Elias *et al.*, 2008). NMFs like trans-urocanic acid and pyrrolidone carboxylic acid are also suggested to play a role in balancing the acidic skin pH. Acidic pH besides regulating the desquamation process, helps to maintain a healthy microbial flora of the skin (Bibel *et al.*, 1987; Rebell *et al.*, 1950).

Of the many mutations reported, two most common loss-of-function mutations (R501X and 2282del4) account for the majority of AD cases in European populations. In Japanese population as well, *FLG* has been shown associated with AD, however, a different profile of *FLG* markers seems to exist in different populations (Osawa *et al.*, 2010). FLG's wider variant landscape was further highlighted with an exhaustive discovery search in the Singapore Chinese population where 22 null mutations, 14 of which were novel, were discovered (Chen *et al.*, 2011). Chen *et al* found different combinations of

these 22 null mutations associated with AD in the Singapore Chinese population ascertained from the National Skin Centre.

A collaborative project was initiated to check the association of the 6 common *FLG* mutations which represented 72% of the *FLG* mutation profile in Singapore.

5.2.2 Materials and method

Samples from the NUS cohort were assayed for the 6 common *FLG* mutants. A total of 700 Singapore Chinese samples were typed. 5 sample failed QC and the remaining 662 AD cases and 33 controls were assayed by Huijia Chen at IMB according to their standardised methods (Chen *et al.*, 2011). The few number of controls were increased by adding the 440 control samples data presented in the Chen *et al* paper. Cases were further split into severity and chronicity and analysed versus controls. Fishers exact test was used to test for associations.

5.2.3 Results

The six common *FLG* mutations found in Singapore were kindly assayed by IMB. Due to the limitation in the number of samples that could be processed by them, we sent all the cases collected as of then and made up the remaining samples with controls.

FLG mutation	% of total *FLG* mutations (paper)	NUS cases	NUS +Paper controls	P value		Paper		
				General	Multiplicative	Cases	Control	P value
c.3321delA	24%	0/25/637	0/11/462	0.229	0.173	0/25/398	0/10/430	0.009
p.S1515X	11%	0/4/658	0/2/471	1.000	0.678	0/10/409	0/2/438	0.019
p.Q2417X	7%	0/5/657	0/0/473	0.079	0.058	0/7/410	0/0/440	0.006
p.E2422X	**6%**	**0/7/655**	**0/0/473**	**0.046**	**0.025**	0/6/416	0/0/440	0.014
p.S2706X	9%	1/8/653	0/8/463	0.773	0.811	0/9/406	0/7/431	0.618
c.6950del8	15%	0/10/652	0/4/469	0.418	0.319	1/14/407	0/4/436	0.011

Table 5-2: Case control association for *FLG* mutations.
Paper = Chen et al, 2011 paper

To conduct a case-control association study, control numbers from the Chen *et al* paper were used, as their population was also Singapore Chinese. Of the 6 mutations assayed, 3 were newly discovered for the paper (Chen *et al.*, 2011). All of the 6 mutations except p.S2706X were found associated with AD in the National Skin Centre cohort. However, only 1 previously known mutation, p.E2422X was found associated in the NUS cohort.

FLG mutation	389 mild AD vs 473 control	189 moderate AD vs 473 control	84 severe AD vs 473 control	242 chronic AD vs 473 control	410 nonchronic AD vs 473 control
c.3321delA	0.529	0.129	0.261	**0.029**	0.830
p.S1515X	0.418	1.000	1.000	0.343	1.000
p.Q2417X	0.092	0.081	1.000	**0.038**	0.215
p.E2422X	**0.041**	**0.023**	1.000	**0.013**	0.100
p.S2706X	0.387	0.732	0.653	0.758	0.490
c.6950del8	0.739	0.234	0.560	0.694	0.527

Table 5-3: Association based on severity and chronicity.

To test the lack of mutation at other positions, cases were divided based on severity and chronicity. The division criteria was identical to section 3.2.2. The results for the association are tabulated above (Table 5.3). pE2422X associated at the case control level of analyses (*P* value= 0.046) was found to

be significantly associated in mild (P =0.041) and in moderate (P =0.023) cases versus controls. This null mutation was also found associated in chronic AD cases.

Two other mutations (c.3321delA, p.Q2417X) were also found associated with chronic AD cases when compared to controls.

5.2.4 Discussion

Of the six tested mutations, only 1 (pE2422X) could be replicated in our cohort. The association for pE2422X showed an increasing trend of association from mild to moderate cases. However, no association to severe cases from NUS was found. This may have been due to the small sample size of 84 severe AD cases versus 473 controls.

The differences in association observed between Chen *et al* study and NUS cohort may be attributed to the different ascertainment procedure. Samples from NSC (Chen *et al.*, 2011) were from clinics where the people were afflicted enough at the time to see the doctor. They also had other dermatological conditions such as ichthyosis vulgaris, keratosis pilaris and a very high percentage of severe AD (40%) and a low mild AD (6%) (Chen *et al.*, 2011). Whereas our cohort was collected at the population level with a very high rate of mild and non-chronic AD. This point is further corroborated with the emergence of associations at c.3321delA, p.Q2417X in chronic cases.

5.2.5 Conclusion

FLG one of the most widely associated gene to AD and ichthyosis vulgaris, is associated with more severe and persistent forms of AD. Milder forms of AD

collected at a population level may have different genes at play in their pathogenesis as shown in Chapter 3.

Chapter 6

Conclusions and Future work

"If you can't explain it simply,
you don't understand it well enough."
-Albert Einstein

6.1 Conclusions

Atopic dermatitis, commonly known as eczema, is a chronic itchy skin condition with an increased incidence of occurrence at skin fold. It is a highly complex disease with a strong genetic and environment component in its pathogenesis. Changing environment associated with industrialisation is positively correlated to AD resulting in a diversity of its prevalence across the globe. With passage of time, different hypotheses for its pathogenesis have gained acceptance. Atopy, described as an IgE-antibody high responder to innocuous agents, was earlier thought to be *the* causative mechanism. Newer hypothesis of a dysfunctional barrier started gaining acceptance in the mid 1990's. This lead to the wider acceptance of the combination of atopy and barrier property related pathways with a dysfunctional barrier as the primary step. Chapter 1 describes the different views pertaining to this ideology and briefly explains the variety of ways the skin maintains homeostasis and its barrier properties, thus, laying the background for subsequent chapters.

In chapter 2 an epidemiological study for eczema in Singapore is explained. This study was done to evaluate the changes since the last 2 published studies.

The comparisons show an increase in prevalence for eczema over the years attributable to strong non-genetic factors (Goh *et al.*, 1996; Tay *et al.*, 2002; Wang *et al.*, 2004). Other parameters of the questionnaire were tested to assess for associations and possible risk factors. Factors such as gender, race and age showed positive association to AD and thus other variables were evaluated after controlling for their confounding effect. Paternal eczema was found to be the strongest association followed by maternal eczema and family history of atopic conditions-- clearly highlighting that AD is strongly linked to genetics and atopy is involved in its pathogenesis.

With the knowledge of genetics as a major contributing factor, we set out to identify barrier-related genes associated with AD in ethnic Chinese. The population ascertained for the epidemiology provided a good sample source for a case-control association study. Direct or indirect barrier related genes were screened and validated through a 2-stage study designed to identify the genes and their risk conferring SNPs. To better discern the results, analyses for chronicity, severity, against SPT positives and controlling for confounding demographic stratifications were conducted. The results showed 27 genes associated with AD of them most were directly linked to barrier maintenance and the remaining through either atopy or autoallergy as described in Chapter 3.

After finding a set of barrier related genes associated with AD, we decided to further characterise the risk conferring SNPs. For the same, *CA6* was selected for further characterisation as described in detail in chapter 4. Eight tags were found associated with the minor allele of 2 SNPs conferring a protective role and that of the remaining 6 conferring risk. Associations of genotypes of these

SNPs to gene expression levels and specific probes for exons were checked independent of disease status. Associations to gene expression and splicing were observed which showed increased gene expression in the risk genotype. However, translated CA6 protein level which is downstream of gene expression and alternative splicing did not show any difference between cases and controls. To evaluate if the SNPs are involved in epigenetic regulation, methylation study was conducted. SNPs were found to be associated with differential methylation pattern when segregated according to genotypes. Increased methylation was found upstream of the gene which is indicative of gene silencing (Keshet *et al.,* 1985). Thus risk genotypes of different SNPs exerted contradictory effects on gene expression which may have been the reason for the lack of CA6 protein level difference between cases and controls. Differential methylation was also observed within the gene when segregated according to SNP genotypes. These SNPs were present in the intron-exon boundaries and may affect the alternative splicing of *CA6*. It has been recently shown that methylation within the gene can affect splicing by either promoting exon inclusion (Maunakea *et al.,* 2013) or exclusion (Shukla *et al.,* 2011). Thus, the associated SNPs of *CA6* were linked to either gene expression and alternative splicing through differential methylation. As *CA6* is hypothesized to affect the pH which in-turn maintains the barrier homeostasis, relevant parameter such as skin pH was measured. Risk conferring genotype of a SNP was found associated with increased pH. Increase in pH has been previously associated with AD (Eberlein-Konig *et al.,* 2000). Increased pH has been shown to be the cause (Holleran *et al.,* 1993) and the effect (Dikstein and Zlotogorski, 1994; Ohman and Vahlquist, 1994) of barrier dysfunction thereby

strengthening the association. As all these associations were at the *ex vivo* level, *in vitro* experiments containing the prevalent haplotypes were initiated. Although, no difference between haplotypes were observed for the luciferase-reporter promoter assay, minigene experiment for checking splicing needs further work.

This thesis identified and validated few novel and many previously associated candidates to AD which can be pursued for functional characterisation. *CA6,* a gene never studied in the context of skin before, is an interesting discovery as it is hypothesized to regulate pH which is the master controller of all barrier processes in the epidermis. Also, this gene has 8 independent signals of association. Functional characterisation of the risk variants to change in gene-function and then ultimately to its involvement in disease pathogenesis may make this the next big candidate after *FLG*.

Chapter 5 gives a brief account of the replication studies conducted on our cohort. The Chinese GWAS SNPs were validated on our population and hints of a possible socio-economic confounder emerged from this work which resulted in a paper (Andiappan *et al.*, 2012). Chen *et al* discovered a different spectrum of *FLG* mutations in the Singapore population and found its association to AD in cases ascertained from National Skin Centre (Chen *et al.*, 2011). Although the common variants were not significantly associated with AD (except one) in our cohort, stratifying our samples based on chronicity and severity showed newer and increased trends of association. This suggests that *FLG* is associated with the more severe and chronic forms of AD and is probably not one of the best descriptors for milder forms of AD.

This thesis has highlighted the importance of other barrier maintaining genes besides *FLG* which are involved in pathogenesis of AD. Knowledge of this can be used to probably make a diagnostic chip of the significant SNPs. Also, the genes found associated will pave way for future research into their respective pathogenic mechanisms. In future, this might form the basis for therapeutic research and formulations.

6.2 Future work and implications

Twenty seven genes identified and validated to be associated with AD can form 27 different major projects of future research. In the context of *CA6,* collaborative efforts to replicate associations in different population would be important to further establish *CA6* as an important AD candidate.

A larger cohort with pH and TEWL measurements would need to be ascertained to increase the representation of the homozygous minor genotypes thereby making the analysis more robust.

Results obtained from different array data-sets would need to be validated by using another platform/technique. Methylation array results would need to be validated on a larger population and by using a different platform. QIAGEN offers assays for validating HumanMethlation450K BeadChip data by pyrosequencing. Alternatively, sequencing after bisulphite conversion can also be adopted (Frommer *et al.*, 1992). Results obtained from the exome array would have to be validated using real-time PCR. *In vitro* splicing studies using mini-genes (Gaildrat *et al.*, 2010) would help to corroborate the exome and gene expression results from the array. Splicing experiments done in presence of; 5-aza-2'-deoxycytidine, a DNA methylation inhibitor (Christman,

2002) and methotrexate, methylation enhancer (Dinh *et al.*, 2013) would help identify the exons included and or excluded due to the risk conferring haplotypes. This would help identify the alternatively spliced variants enriched through different haplotypes under the regulatory effect of methylation.

To carefully tease out the effects of the selectively spliced forms, stable cell lines expressing these variants could be constructed. Measurements such as total expressed CA6 protein could be measured to validate the effect on expression levels. Further physiologically relevant experiments such as wound healing (Liang *et al.*, 2007) could be conducted on these stable cell lines. Measurements such as trans-epithelial electrical resistance, a measure of permeability of the monolayer could be conducted (Ebihara *et al.*, 2005). Alternatively, a labelled solute flux assay could be used to again test for the integrity of the monolayer containing the risk variants (Chandra *et al.*, 2007).

The stable cell lines could also be used to detect a change in pH by using different adapter to the skin pH meter. If a difference in pH and monolayer integrity is observed then the intermediate pathways amounting to the observation could be evaluated. Over expressing and silenced cells for CA6 could be analysed for gene expression and different genes up or down-regulated could be then examined for their role in pathway leading to barrier defect and or change in pH.

Also, organotypic models can be created for the CA6 variants for future basic and therapeutic research (Oh *et al.*, 2013). Our collaborators at IMB have the expertise in constructing 3D skin models which can be used for a whole set of experiments for cross-validation and also becoming a skin-model for future

therapeutic research.

List of Publications

1. ‘Genome-wide association study for atopy and allergic rhinitis in Singapore Chinese population’. Anand Kumar Andiappan, De Yun Wang, Ramani Anantharaman, Pallavi Nilkanth Parate, **Bani Kaur Suri**, Hui Qi Low, Yi Li, Wanting Zhao, Paola Castagnoli, Jianjun Liu, Fook Tim Chew. *PLoS One 2011*. 6(5): e19719. doi:10.1371/journal.pone.0019719

2. ‘Variation in Uteroglobin-Related Protein 1 (UGRP1) gene is associated with Allergic Rhinitis in Singapore Chinese’. Anand Kumar Andiappan, Wei Sheng Yeo, Pallavi Nilkanth Parate, Ramani Anantharaman, **Bani Kaur Suri**, Yun De Wang, Fook Tim Chew,2011. *BMC Medical Genetics 2011*, 12:1 doi: 10.1186/1471-2350-12-39

3. ‘Genetic variation in BDNF is associated with allergic asthma and allergic rhinitis in an ethnic Chinese population in Singapore’. Anand Kumar Andiappan, Pallavi Nilkanth Parate, Ramani Anantharaman, **Bani Kaur Suri**, De Yun Wang, Fook Tim Chew. *Cytokine 2011*. 56:2 doi: 10.1016/j.cyto.2011.05.008

4. ‘Genome-wide association study identifies PERLD1 as asthma candidate gene’. Ramani Anantharaman, Anand K Andiappan, Pallavi P Nilkanth, **Bani K Suri**, De Yun Wang and Fook Tim Chew. *BMC Medical Genetics 2011*. 12:170 doi:10.1186/1471-2350-12-170

5. ‘Validation of GWAS Loci for Atopic Dermatitis in a Singapore Chinese Population’. Anand K Andiappan, Jia N Foo, Meng W Choy, Huijia Chen, John E A Common, Mark B Y Tang, Hugo P van Bever, Yoke C Giam, **Bani K Suri**, Anantharaman Ramani, Parate P Nilkanth, E Birgitte Lane, De Y Wang, Fook T Chew and Jianjun Liu. *Journal of Investigative Dermatology 2012*. 132 doi:10.1038/jid.2011.471

Bibliography

Akdis CA, Akdis M, Bieber T, Bindslev-Jensen C, Boguniewicz M, Eigenmann P, *et al.* (2006) Diagnosis and treatment of atopic dermatitis in children and adults: European Academy of Allergology and Clinical Immunology/American Academy of Allergy, Asthma and Immunology/PRACTALL Consensus Report. *J Allergy Clin Immunol* 118:152-69.

Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, *et al.* (2010) Integrating common and rare genetic variation in diverse human populations. *Nature* 467:52-8.

Andiappan AK, Anantharaman R, Nilkanth PP, Wang de Y, Chew FT (2010) Evaluating the transferability of Hapmap SNPs to a Singapore Chinese population. *BMC genetics* 11:36.

Andiappan AK, Foo JN, Choy MW, Chen H, Common JE, Tang MB, *et al.* (2012) Validation of GWAS loci for atopic dermatitis in a Singapore Chinese population. *The Journal of investigative dermatology* 132:1505-7.

Angelova-Fischer I, Bauer A, Hipler UC, Petrov I, Kazandjieva J, Bruckner T, *et al.* (2005) The Objective Severity Assessment of Atopic Dermatitis (OSAAD) score: validity, reliability and sensitivity in adult patients with atopic dermatitis. *British Journal of Dermatology* 153:767-73.

Ardlie KG, Kruglyak L, Seielstad M (2002) Patterns of linkage disequilibrium in the human genome. *Nature reviews Genetics* 3:299-309.

Arikawa J, Ishibashi M, Kawashima M, Takagi Y, Ichikawa Y, Imokawa G (2002) Decreased levels of sphingosine, a natural antimicrobial agent, may be associated with vulnerability of the stratum corneum from patients with atopic dermatitis to colonization by Staphylococcus aureus. *The Journal of investigative dermatology* 119:433-9.

Asher MI, Keil U, Anderson HR, Beasley R, Crane J, Martinez F, *et al.* (1995) International study of asthma and allergies in childhood (ISAAC): rationale and methods. *European Respiratory Journal* 8:483-91.

Baek JO, Hong S, Son DK, Lee JR, Roh JY, Kwon HJ (2013) Analysis of the prevalence of and risk factors for atopic dermatitis using an ISAAC questionnaire in 8,750 Korean children. *Int Arch Allergy Immunol* 162:79-85.

Barker JN, Palmer CN, Zhao Y, Liao H, Hull PR, Lee SP, *et al.* (2007) Null mutations in the filaggrin gene (FLG) determine major susceptibility to early-onset atopic dermatitis that persists into adulthood. *The Journal of investigative dermatology* 127:564-7.

Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, Rioux JD, *et al.* (2008) Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nature genetics* 40:955-62.

Bashir SJ, Chew AL, Anigbogu A, Dreher F, Maibach HI (2001) Physical and physiological effects of stratum corneum tape stripping. *Skin research and technology : official journal of International Society for Bioengineering and the Skin (ISBS) [and] International Society for Digital Imaging of Skin (ISDIS) [and] International Society for Skin Imaging (ISSI)* 7:40-8.

Batlles Garrido J, Torres-Borrego J, Bonillo Perales A, Rubi Ruiz T, Gonzalez Jimenez Y, Momblan De Cabo J, *et al.* (2010) Prevalence and factors linked to atopic eczema in 10- and 11-year-old schoolchildren. Isaac 2 in Almeria, Spain. *Allergologia et immunopathologia* 38:174-80.

Bibel DJ, Aly R, Lahti L, Shinefield HR, Maibach HI (1987) Microbial adherence to vulvar epithelial cells. *Journal of medical microbiology* 23:75-82.

Bieber T (2010) Atopic Dermatitis. *Annals of Dermatology* 22:125.

Bigliardi PL, Tobin DJ, Gaveriaux-Ruff C, Bigliardi-Qi M (2009) Opioids and the skin – where do we stand? *Experimental Dermatology* 18:424-30.

Blencowe BJ (2006) Alternative splicing: new insights from global analyses. *Cell* 126:37-47.

Boerwinkle E, Chakraborty R, Sing CF (1986) The use of measured genotype information in the analysis of quantitative phenotypes in man. I. Models and analytical methods. *Annals of human genetics* 50:181-94.

Borecki IB, Province MA (2008) Linkage and Association: Basic Concepts. In: *Advances in Genetics* (Rao DC, Gu CC, eds) Vol. Volume 60: Academic Press, 51-74.

Boukamp P, Petrussevska RT, Breitkreutz D, Hornung J, Markham A, Fusenig NE (1988) Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. *The Journal of cell biology* 106:761-71.

Boyes J, Bird A (1991) DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* 64:1123-34.

Brattsand M, Stefansson K, Lundh C, Haasum Y, Egelrud T (2005) A proteolytic cascade of kallikreins in the stratum corneum. *The Journal of investigative dermatology* 124:198-203.

Braun-Falco O, Korting HC (1986) Der normale pH-Wert der menschlichen Haut. *Der Hautarzt* 37:4.

Breuckmann F, von Kobyletzki G, Avermaete A, Kreuter A, Altmeyer P, Gambichler T (2002) Modulation of cathepsin G expression in severe atopic dermatitis following medium-dose UVA1 phototherapy. *BMC Dermatol* 2:1-5.

Broberg A, Svensson A, Borres MP, Berg R (2000) Atopic dermatitis in 5-6-year-old Swedish children: cumulative incidence, point prevalence, and severity scoring. *Allergy* 55:1025-9.

Cabral A, Voskamp P, Cleton-Jansen AM, South A, Nizetic D, Backendorf C (2001) Structural organization and regulation of the small proline-rich family of cornified envelope precursors suggest a role in adaptive barrier function. *The Journal of biological chemistry* 276:19231-7.

Candi E, Schmidt R, Melino G (2005) The cornified envelope: a model of cell death in the skin. *Nature reviews Molecular cell biology* 6:328-40.

Cao Y, Liao M, Huang X, Mo Z, Gao F (2009) Meta-analysis of genome-wide linkage studies of atopic dermatitis. *Dermatitis : contact, atopic, occupational, drug* 20:193-9.

Carter MJ (1972) Carbonic anhydrase: isoenzymes, properties, distribution, and functional significance. *Biological reviews of the Cambridge Philosophical Society* 47:465-513.

Chandra A, Barillas S, Suliman A, Angle N (2007) A novel fluorescence-based cellular permeability assay. *Journal of biochemical and biophysical methods* 70:329-33.

Chapman MD, Wood RA (2001) The role and remediation of animal allergens in allergic diseases. *J Allergy Clin Immunol* 107:S414-21.

Chen H, Common JE, Haines RL, Balakrishnan A, Brown SJ, Goh CS, *et al.* (2011) Wide spectrum of filaggrin-null mutations in atopic dermatitis highlights differences between Singaporean Chinese and European populations. *Br J Dermatol* 165:106-14.

Chew FT, Lim SH, Goh DY, Lee BW (1999) Sensitization to local dust-mite fauna in Singapore. *Allergy* 54:1150-9.

Christman JK (2002) 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene* 21:5483-95.

Chupp GL, Lee CG, Jarjour N, Shim YM, Holm CT, He S, *et al.* (2007) A Chitinase-like Protein in the Lung and Circulation of Patients with Severe Asthma. *New England Journal of Medicine* 357:2016-27.

Committee IS (1998) Worldwide variations in the prevalence of asthma symptoms: the International Study of Asthma and Allergies in Childhood (ISAAC). *European Respiratory Journal* 12:315-35.

Common JE, Brown SJ, Haines RL, Goh CS, Chen H, Balakrishnan A, *et al.* (2011) Filaggrin null mutations are not a protective factor for acne vulgaris. *The Journal of investigative dermatology* 131:1378-80.

Cookson WO, Ubhi B, Lawrence R, Abecasis GR, Walley AJ, Cox HE, *et al.* (2001) Genetic linkage of childhood atopic dermatitis to psoriasis susceptibility loci. *Nature genetics* 27:372-3.

Cork M, Robinson D, Vasilopoulos Y, Ferguson A, Moustafa M, Macgowan A, *et al.* (2006) New perspectives on epidermal barrier dysfunction in atopic dermatitis: Gene–environment interactions. *Journal of Allergy and Clinical Immunology* 118:3-21.

Cork MJ, Britton J, Butler L, Young S, Murphy R, Keohane SG (2003) Comparison of parent knowledge, therapy utilization and severity of atopic eczema before and after explanation and demonstration of topical therapies by a specialist dermatology nurse. *Br J Dermatol* 149:582-9.

Cork MJ, Danby SG, Vasilopoulos Y, Hadgraft J, Lane ME, Moustafa M, *et al.* (2009) Epidermal Barrier Dysfunction in Atopic Dermatitis. *Journal of Investigative Dermatology* 129:1892-908.

Costa V, Aprile M, Esposito R, Ciccodicola A (2013) RNA-Seq and human complex diseases: recent accomplishments and future perspectives. *European journal of human genetics : EJHG* 21:134-42.

Coulon V, Chebli K, Cavelier P, Blanchard JM (2010) A novel mouse c-fos intronic promoter that responds to CREB and AP-1 is developmentally regulated in vivo. *PloS one* 5:e11235.

Datta S, Milner JD (2011) Altered T-cell receptor signaling in the pathogenesis of allergic disease. *Journal of Allergy and Clinical Immunology* 127:351-4.

Dawn Teare M, Barrett JH (2005) Genetic linkage studies. *The Lancet* 366:1036-44.

de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D (2005) Efficiency and power in genetic association studies. *Nature genetics* 37:1217-23.

De Benedetto A, Latchney LR, McGirt LY, Vidyasagar S, Cheadle C, Barnes KC, *et al.* (2008) The Tight Junction Protein, Claudin-1 is Dysregulated in Atopic Dermatitis. *The Journal of allergy and clinical immunology* 121:S32.

De Benedetto A, Rafaels NM, McGirt LY, Ivanov AI, Georas SN, Cheadle C, *et al.* (2011) Tight junction defects in patients with atopic dermatitis. *J Allergy Clin Immunol* 127:773-86 e1-7.

Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P, *et al.* (2002) Comprehensive human genome amplification using multiple displacement amplification. *Proceedings of the National Academy of Sciences of the United States of America* 99:5261-6.

Delea TE, Gokhale M, Makin C, Hussein MA, Vanderpoel J, Sandman T, *et al.* (2007) Administrative claims analysis of utilization and costs of care in health plan members with atopic dermatitis who had prior use of a topical corticosteroid and who initiate therapy with pimecrolimus or tacrolimus. *J Manag Care Pharm* 13:349-59.

Diepgen TL, Fartasch M (1992) Recent epidemiological and genetic studies in atopic dermatitis. *Acta dermato-venereologica Supplementum* 176:13-8.

Dikstein S, Zlotogorski A (1994) Measurement of skin pH. *Acta dermato-venereologica Supplementum* 185:18-20.

Dinh TT, O'Leary M, Won SY, Li S, Arroyo L, Liu X, *et al.* (2013) Generation of a luciferase-based reporter for CHH and CG DNA methylation in Arabidopsis thaliana. *Silence* 4:1.

Dreher F, Modjtahedi BS, Modjtahedi SP, Maibach HI (2005) Quantification of stratum corneum removal by adhesive tape stripping by total protein assay in 96-well microplates. *Skin research and technology : official journal of International Society for Bioengineering and the Skin (ISBS) [and] International Society for Digital Imaging of Skin (ISDIS) [and] International Society for Skin Imaging (ISSI)* 11:97-101.

Eberlein-Konig B, Schafer T, Huss-Marp J, Darsow U, Mohrenschlager M, Herbert O, *et al.* (2000) Skin surface pH, stratum corneum hydration, trans-epidermal water loss and skin roughness related to atopic eczema and skin dryness in a population of primary school children. *Acta Derm Venereol* 80:188-91.

Ebihara N, Funaki T, Murakami A, Takai S, Miyazaki M (2005) Mast cell chymase decreases the barrier function and inhibits the migration of corneal epithelial cells. *Current eye research* 30:1061-9.

Eden S, Cedar H (1994) Role of DNA methylation in the regulation of transcription. *Current opinion in genetics & development* 4:255-9.

Ekholm IE, Brattsand M, Egelrud T (2000) Stratum corneum tryptic enzyme in normal epidermis: a missing link in the desquamation process? *The Journal of investigative dermatology* 114:56-63.

Elias P, Hatano Y, Williams M (2008) Basis for the barrier abnormality in atopic dermatitis: Outside-inside-outside pathogenic mechanisms. *Journal of Allergy and Clinical Immunology* 121:1337-43.

Elias PM (1983) Epidermal lipids, barrier function, and desquamation. *The Journal of investigative dermatology* 80:44s-9s.

Elias PM, Schmuth M (2009) Abnormal skin barrier in the etiopathogenesis of atopic dermatitis. *Current Opinion in Allergy and Clinical Immunology* 9:437-46.

Emami N, Diamandis EP (2008) Human kallikrein-related peptidase 14 (KLK14) is a new activator component of the KLK proteolytic cascade. Possible function in seminal plasma and skin. *The Journal of biological chemistry* 283:3031-41.

Esparza-Gordillo J, Weidinger S, Fölster-Holst R, Bauerfeind A, Ruschendorf F, Patone G, *et al.* (2009) A common variant on chromosome 11q13 is associated with atopic dermatitis. *Nature genetics* 41:596-601.

Fallon PG, Sasaki T, Sandilands A, Campbell LE, Saunders SP, Mangan NE, *et al.* (2009) A homozygous frameshift mutation in the mouse Flg gene facilitates enhanced percutaneous allergen priming. *Nature Genetics* 41:602-8.

Fergusson DM, Horwood LJ, Shannon FT (1982) Risk factors in childhood eczema. *Journal of epidemiology and community health* 36:118-22.

Fernley RT, Wright RD, Coghlan JP (1979) A novel carbonic anhydrase from the ovine parotid gland. *FEBS letters* 105:299-302.

Flohr C, Johansson SGO, Wahlgren C-F, Williams H (2004) How atopic is atopic dermatitis? *Journal of Allergy and Clinical Immunology* 114:150-8.

Fluhr JW, Elias PM (2002) Stratum corneum pH: Formation and Function of the ‘Acid Mantle&rsquo. *Exogenous Dermatology* 1:163-75.

Fluhr JW, Feingold KR, Elias PM (2006) Transepidermal water loss reflects permeability barrier status: validation in human and rodent in vivo and ex vivo models. *Exp Dermatol* 15:483-92.

Fok AOL, Wong GWK (2009) What have we learnt from ISAAC phase III in the Asia-Pacific rim? *Current Opinion in Allergy & Clinical Immunology* 9:7.

Frateschi S, Camerer E, Crisante G, Rieser S, Membrez M, Charles RP, *et al.* (2011) PAR2 absence completely rescues inflammation and ichthyosis caused by altered CAP1/Prss8 expression in mouse skin. *Nature communications* 2:161.

Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, *et al.* (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences of the United States of America* 89:1827-31.

Gaildrat P, Killian A, Martins A, Tournier I, Frebourg T, Tosi M (2010) Use of splicing reporter minigene assay to evaluate the effect on splicing of unclassified genetic variants. *Methods in molecular biology (Clifton, NJ)* 653:249-57.

Galli SJ, Tsai M, Piliponsky AM (2008) The development of allergic inflammation. *Nature* 454:445-54.

Garrett JP, Apter AJ, Hoffstad O, Spergel JM, Margolis DJ (2013) Asthma and frequency of wheeze: risk factors for the persistence of atopic dermatitis in children. *Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology* 110:146-9.

Gasser T, Muller-Myhsok B, Wszolek ZK, Oehlmann R, Calne DB, Bonifati V, *et al.* (1998) A susceptibility locus for Parkinson's disease maps to chromosome 2p13. *Nature genetics* 18:262-5.

Glaser AG, Menz G, Kirsch AI, Zeller S, Crameri R, Rhyner C (2008) Auto- and cross-reactivity to thioredoxin allergens in allergic bronchopulmonary aspergillosis. *Allergy* 63:1617-23.

Goh DY, Chew FT, Quek SC, Lee BW (1996) Prevalence and severity of asthma, rhinitis, and eczema in Singapore schoolchildren. *Arch Dis Child* 74:131-5.

Gonda A, Gal I, Szanto S, Sarraj B, Glant TT, Hunyadi J, *et al.* (2005) CD44, but not l-selectin, is critically involved in leucocyte migration into the skin in a murine model of allergic dermatitis. *Exp Dermatol* 14:700-8.

Goodwin H (2008) Eczema and allergy: how useful is allergy testing? *Paediatric nursing* 20:25-30.

Gudjonsson JE, Ding J, Johnston A, Tejasvi T, Guzman AM, Nair RP, *et al.* (2010) Assessment of the psoriatic transcriptome in a large sample: additional regulated genes and comparisons with in vitro models. *The Journal of investigative dermatology* 130:1829-40.

Guttman-Yassky E, Lowes MA, Fuentes-Duculan J, Zaba LC, Cardinale I, Nograles KE, *et al.* (2008) Low Expression of the IL-23/Th17 Pathway in Atopic Dermatitis Compared to Psoriasis. *The Journal of Immunology* 181:7420-7.

Guttman-Yassky E, Nograles KE, Krueger JG (2011) Contrasting pathogenesis of atopic dermatitis and psoriasis--part I: clinical and pathologic concepts. *J Allergy Clin Immunol* 127:1110-8.

Hachem JP, Crumrine D, Fluhr J, Brown BE, Feingold KR, Elias PM (2003) pH directly regulates epidermal permeability barrier homeostasis, and stratum corneum integrity/cohesion. *The Journal of investigative dermatology* 121:345-53.

Hachem JP, Wagberg F, Schmuth M, Crumrine D, Lissens W, Jayakumar A, *et al.* (2006) Serine protease activity and residual LEKTI expression determine phenotype in Netherton syndrome. *The Journal of investigative dermatology* 126:1609-21.

Han J, Xiong J, Wang D, Fu XD (2011) Pre-mRNA splicing: where and when in the nucleus. *Trends in cell biology* 21:336-43.

Hanakawa Y, Schechter NM, Lin C, Garza L, Li H, Yamaguchi T, *et al.* (2002) Molecular mechanisms of blister formation in bullous impetigo and staphylococcal scalded skin syndrome. *The Journal of clinical investigation* 110:53-60.

Hanifin JM, Rajka G (1980) Diagnostic features of atopic dermatitis. *Acta Dermato-Venereologica (Stockholm)*:4.

Hanifin JM, Reed ML (2007) A population-based survey of eczema prevalence in the United States. *Dermatitis : contact, atopic, occupational, drug* 18:82-91.

Harder J, Dressel S, Wittersheim M, Cordes J, Meyer-Hoffert U, Mrowietz U, *et al.* (2010) Enhanced expression and secretion of antimicrobial peptides in atopic dermatitis and after superficial skin injury. *The Journal of investigative dermatology* 130:1355-64.

Hata T, Gallo R (2008) Antimicrobial Peptides, Skin Infections, and Atopic Dermatitis. *Seminars in Cutaneous Medicine and Surgery* 27:144-50.

Hetjss E (1892) Die Reaktion des Schweisses beim gesunden Menschen. *Monatsh f prakt Dermat* 14:17.

Hirota T, Takahashi A, Kubo M, Tsunoda T, Tomita K, Sakashita M, *et al.* (2012) Genome-wide association study identifies eight new susceptibility loci for atopic dermatitis in the Japanese population. *Nature genetics* 44:1222-6.

Holgate ST, Roberts G, Arshad HS, Howarth PH, Davies DE (2009) The Role of the Airway Epithelium and its Interaction with Environmental Factors in Asthma Pathogenesis. *Proceedings of the American Thoracic Society* 6:655-9.

Holleran WM, Takagi Y, Menon GK, Legler G, Feingold KR, Elias PM (1993) Processing of epidermal glucosylceramides is required for optimal mammalian cutaneous permeability barrier function. *The Journal of clinical investigation* 91:1656-64.

Hong SH, Nah HY, Lee JY, Lee YJ, Lee JW, Gye MC, *et al.* (2004) Estrogen regulates the expression of the small proline-rich 2 gene family in the mouse uterus. *Molecules and cells* 17:477-84.

Hosono S, Faruqi AF, Dean FB, Du Y, Sun Z, Wu X, *et al.* (2003) Unbiased whole-genome amplification directly from clinical samples. *Genome research* 13:954-64.

Hughes-Davies L, Huntsman D, Ruas M, Fuks F, Bye J, Chin SF, *et al.* (2003) EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. *Cell* 115:523-35.

Illi S, von Mutius E, Lau S, Nickel R, Gruber C, Niggemann B, *et al.* (2004) The natural course of atopic dermatitis from birth to age 7 years and the association with asthma. *J Allergy Clin Immunol* 113:925-31.

Imokawa G, Abe A, Jin K, Higaki Y, Kawashima M, Hidano A (1991) Decreased level of ceramides in stratum corneum of atopic dermatitis: an etiologic factor in atopic dry skin? *J Invest Dermatol* 96:523-6.

ISAAC (1993) <Phase 1 manual.pdf>.

Jaafar RB, Pettit JH (1993) Atopic eczema in a multiracial country (Malaysia). *Clinical and experimental dermatology* 18:496-9.

Jarzab J, Filipowska B, Zebracka J, Kowalska M, lstrok, gorzata, *et al.* (2010) Locus 1q21 Gene Expression Changes in Atopic Dermatitis Skin Lesions: Deregulation of Small Proline-Rich Region 1A. *International Archives of Allergy and Immunology* 151:28-37.

Johansson SGO, Bieber T, Dahl R, Friedmann PS, Lanier BQ, Lockey RF, *et al.* (2004) Revised nomenclature for allergy for global use: Report of the Nomenclature Review Committee of the World Allergy Organization, October 2003. *Journal of Allergy and Clinical Immunology* 113:832-6.

Kalantari-Dehaghi M, Anhalt GJ, Camilleri MJ, Chernyavsky AI, Chun S, Felgner PL, *et al.* (2013) Pemphigus vulgaris autoantibody profiling by proteomic technique. *PloS one* 8:e57587.

Kamsteeg M, Zeeuwen PL, de Jongh GJ, Rodijk-Olthuis D, Zeeuwen-Franssen ME, van Erp PE, *et al.* (2007) Increased expression of carbonic anhydrase II (CA II) in lesional skin of atopic dermatitis: regulation by Th2 cytokines. *The Journal of investigative dermatology* 127:1786-9.

Karhumaa P, Leinonen J, Parkkila S, Kaunisto K, Tapanainen J, Rajaniemi H (2001) The identification of secreted carbonic anhydrase VI as a constitutive glycoprotein of human and rat milk. *Proceedings of the National Academy of Sciences of the United States of America* 98:11604-8.

Kass SU, Landsberger N, Wolffe AP (1997) DNA methylation directs a time-dependent repression of transcription initiation. *Current biology : CB* 7:157-65.

Keshet I, Yisraeli J, Cedar H (1985) Effect of regional DNA methylation on gene expression. *Proceedings of the National Academy of Sciences of the United States of America* 82:2560-4.

Khalifah RG (1971a) The carbon dioxide hydration activity of carbonic anhydrase. I. Stop-flow kinetic studies on the native human isoenzymes B and C. *The Journal of biological chemistry* 246:2561-73.

Khalifah RG (1971b) The Carbon Dioxide Hydration Activity of Carbonic Anhydrase: I. STOP-FLOW KINETIC STUDIES ON THE NATIVE HUMAN ISOENZYMES B AND C. *Journal of Biological Chemistry* 246:2561-73.

Kim N, Bae KB, Kim MO, Yu DH, Kim HJ, Yuh HS, *et al.* (2012) Overexpression of cathepsin S induces chronic atopic dermatitis in mice. *The Journal of investigative dermatology* 132:1169-76.

Kluken H, Wienker T, Bieber T (2003) Atopic eczema/dermatitis syndrome - a genetically complex disease. New advances in discovering the genetic contribution. *Allergy* 58:5-12.

Kniep EM, Roehlecke C, Ozkucur N, Steinberg A, Reber F, Knels L, *et al.* (2006) Inhibition of apoptosis and reduction of intracellular pH decrease in retinal neural cell cultures by a blocker of carbonic anhydrase. *Investigative ophthalmology & visual science* 47:1185-92.

Komatsu N, Saijoh K, Kuk C, Liu AC, Khan S, Shirasaki F, *et al.* (2007) Human tissue kallikrein expression in the stratum corneum and serum of atopic dermatitis patients. *Experimental Dermatology* 16:513-9.

Kraft S, Wessendorf JH, Hanau D, Bieber T (1998) Regulation of the high affinity receptor for IgE on human epidermal Langerhans cells. *J Immunol* 161:1000-6.

Krawczak M, Thomas NS, Hundrieser B, Mort M, Wittig M, Hampe J, *et al.* (2007) Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing. *Human mutation* 28:150-8.

Kruglyak L, Nickerson DA (2001) Variation is the spice of life. *Nature genetics* 27:234-6.

Lanier LL, O'Fallon S, Somoza C, Phillips JH, Linsley PS, Okumura K, *et al.* (1995) CD80 (B7) and CD86 (B70) provide similar costimulatory signals for T cell proliferation, cytokine production, and generation of CTL. *J Immunol* 154:97-105.

Larsen FS, Diepgen T, Svensson Å (1996) The occurrence of atopic dermatitis in North Europe: An international questionnaire study. *Journal of the American Academy of Dermatology* 34:760-4.

Lau YL, Karlberg J, Yeung CY (1995) Prevalence of and factors associated with childhood asthma in Hong Kong. *Acta paediatrica (Oslo, Norway : 1992)* 84:820-2.

Lee Y, Hwang K (2002) Skin thickness of Korean adults. *Surg Radiol Anat* 24:183-9.

Lee YA, Wahn U, Kehrt R, Tarani L, Businco L, Gustafsson D, *et al.* (2000) A major susceptibility locus for atopic dermatitis maps to chromosome 3q21. *Nature genetics* 26:470-3.

Lewis CM (2002) Genetic association studies: design, analysis and interpretation. *Briefings in bioinformatics* 3:146-53.

Lewis CM, Knight J (2012) Introduction to Genetic Association Studies. *Cold Spring Harbor Protocols* 2012:pdb.top068163.

Leyvraz C, Charles RP, Rubera I, Guitard M, Rotman S, Breiden B, *et al.* (2005) The epidermal barrier function is dependent on the serine protease CAP1/Prss8. *The Journal of cell biology* 170:487-96.

Liang CC, Park AY, Guan JL (2007) In vitro scratch assay: a convenient and inexpensive method for analysis of cell migration in vitro. *Nature protocols* 2:329-33.

Loden M (1995) Biophysical properties of dry atopic and normal skin with special reference to effects of skin care products. *Acta dermato-venereologica Supplementum* 192:1-48.

Lorincz MC, Dickerson DR, Schmitt M, Groudine M (2004) Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nature structural & molecular biology* 11:1068-75.

Man M, Elias PM, Man W, Wu Y, Bourguignon LYW, Feingold KR, *et al.* (2009) The role of CD44 in cutaneous inflammation. *Experimental Dermatology* 18:962-8.

Marchionini A, Hausknecht W (1938) SaÈ uremantel der Haut und Bakterienabwehr. *Klin Wochenschr* 17:4.

Margolis DJ, Gupta J, Apter AJ, Ganguly T, Hoffstad O, Papadopoulos M, *et al.* (2013) Filaggrin-2 variation is associated with more persistent atopic dermatitis in African American subjects. *J Allergy Clin Immunol*.

Martin PE, Koplin JJ, Eckert JK, Lowe AJ, Ponsonby AL, Osborne NJ, *et al.* (2013) The prevalence and socio-demographic risk factors of clinical eczema in infancy: a population-based observational study. *Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology* 43:642-51.

Matlin AJ, Clark F, Smith CW (2005) Understanding alternative splicing: towards a cellular code. *Nature reviews Molecular cell biology* 6:386-98.

Maunakea AK, Chepelev I, Cui K, Zhao K (2013) Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell research* 23:1256-69.

Mauro T, Holleran WM, Grayson S, Gao WN, Man MQ, Kriehuber E, *et al.* (1998) Barrier recovery is impeded at neutral pH, independent of ionic effects: implications for extracellular lipid processing. *Archives of dermatological research* 290:215-22.

Miyake Y, Tanaka K, Sasaki S, Hirota Y (2012) Parental employment, income, education and allergic disorders in children: a prebirth cohort study in Japan. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 16:756-61.

Morales Suarez-Varela M, Garcia-Marcos L, Kogan MD, Llopis Gonzalez A, Martinez Gimeno A, Aguinaga Ontoso I, *et al.* (2008) Parents' smoking habit and prevalence of atopic eczema in 6-7 and 13-14 year-old schoolchildren in Spain. ISAAC phase III. *Allergologia et immunopathologia* 36:336-42.

Nakamura Y (2009) DNA variations in human and medical genetics: 25 years of my experience. *Journal of human genetics* 54:1-8.

Nakhoul NL, Davis BA, Romero MF, Boron WF (1998) Effect of expressing the water channel aquaporin-1 on the CO2 permeability of Xenopus oocytes. *The American journal of physiology* 274:C543-8.

Nikolovski J, Stamatas GN, Kollias N, Wiegand BC (2008) Barrier function and water-holding and transport properties of infant stratum corneum are different from adult and continue to develop through the first year of life. *The Journal of investigative dermatology* 128:1728-36.

Nilsen TW, Graveley BR (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463:457-63.

Ogawa H, Yoshiike T (1993) A speculative view of atopic dermatitis: barrier dysfunction in pathogenesis. *Journal of dermatological science* 5:197-204.

Ogawa K, Hashida R, Miyagawa M, Kagaya S, Sugita Y, Matsumoto K, *et al.* (2003) Analysis of gene expression in peripheral blood eosinophils from patients with atopic dermatitis and in vitro cytokine-stimulated blood eosinophils. *Clinical and experimental immunology* 131:436-45.

Oh JW, Hsi TC, Guerrero-Juarez CF, Ramos R, Plikus MV (2013) Organotypic skin culture. *The Journal of investigative dermatology* 133:e14.

Ohman H, Vahlquist A (1994) In vivo studies concerning a pH gradient in human stratum corneum and upper epidermis. *Acta Derm Venereol* 74:375-9.

Okoniewski MJ, Hey Y, Pepper SD, Miller CJ (2007) High correspondence between Affymetrix exon and standard expression arrays. *BioTechniques* 42:181-5.

Oranje AP, Glazenburg EJ, Wolkerstorfer A, de Waard-van der Spek FB (2007) Practical issues on interpretation of scoring atopic dermatitis: the SCORAD index, objective SCORAD and the three-item severity score. *Br J Dermatol* 157:645-8.

Oregan G, Sandilands A, McLean W, Irvine A (2008) Filaggrin in atopic dermatitis. *Journal of Allergy and Clinical Immunology* 122:689-93.

Osawa R, Konno S, Akiyama M, Nemoto-Hasebe I, Nomura T, Nomura Y, *et al.* (2010) Japanese-specific filaggrin gene mutations in Japanese patients suffering from atopic eczema and asthma. *J Invest Dermatol* 130:2834-6.

Ozdemir O, Erol AY (2013) Preventative and therapeutic probiotic use in allergic skin conditions: experimental and clinical findings. *BioMed research international* 2013:932391.

Palmer CNA, Irvine AD, Terron-Kwiatkowski A, Zhao Y, Liao H, Lee SP, *et al.* (2006) Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nature genetics* 38:441-6.

Parkkila S, Parkkila AK, Juvonen T, Rajaniemi H (1994) Distribution of the carbonic anhydrase isoenzymes I, II, and VI in the human alimentary tract. *Gut* 35:646-50.

Patel S, Xi ZF, Seo EY, McGaughey D, Segre JA (2006) Klf4 and corticosteroids activate an overlapping set of transcriptional targets to

accelerate in utero epidermal barrier acquisition. *Proceedings of the National Academy of Sciences of the United States of America* 103:18668-73.

Pelucchi C, Galeone C, Bach JF, La Vecchia C, Chatenoud L (2013) Pet exposure and risk of atopic dermatitis at the pediatric age: a meta-analysis of birth cohort studies. *J Allergy Clin Immunol* 132:616-22 e7.

Peres RC, Camargo G, Mofatto LS, Cortellazzi KL, Santos MC, Nobre-dos-Santos M, *et al.* (2010) Association of polymorphisms in the carbonic anhydrase 6 gene with salivary buffer capacity, dental plaque pH, and caries index in children aged 7-9 years. *The pharmacogenomics journal* 10:114-9.

Phillips T (2008) The Role of Methylation in Gene Expression. *Nature Education* 1:3.

Poliani PL, Fontana E, Roifman CM, Notarangelo LD (2013) ζ Chain–associated protein of 70 kDa (ZAP70) deficiency in human subjects is associated with abnormalities of thymic stromal cells: Implications for T-cell tolerance. *The Journal of allergy and clinical immunology* 131:597-600.e2.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* 81:559-75.

Raymond AA, Gonzalez de Peredo A, Stella A, Ishida-Yamamoto A, Bouyssie D, Serre G, *et al.* (2008) Lamellar bodies of human epidermis: proteomics characterization by high throughput mass spectrometry and possible involvement of CLIP-170 in their trafficking/secretion. *Molecular & cellular proteomics : MCP* 7:2151-75.

Rebell G, Pillsbury DM, De Saint Phalle M, Ginsburg D (1950) Factors affecting the rapid disappearance of bacteria placed on the normal skin. *The Journal of investigative dermatology* 14:247-64.

Roifman CM, Dadi H, Somech R, Nahum A, Sharfe N (2010) Characterization of ζ-associated protein, 70 kd (ZAP70)–deficient human lymphocytes. *The Journal of allergy and clinical immunology* 126:1226-33.e1.

Ronchetti R, Haluszka J, Martella S, Falasca C, Guglielmi F, Parmiani S, *et al.* (2003) Skin reactivity to histamine and to allergens in unselected 9-year-old children living in Poland and Italy. *Pediatric allergy and immunology : official publication of the European Society of Pediatric Allergy and Immunology* 14:201-6.

Ross-Hansen K, Linneberg A, Johansen JD, Hersoug LG, Brasch-Andersen C, Menne T, *et al.* (2013) The role of glutathione S-transferase and claudin-1 gene polymorphisms in contact sensitization: a cross-sectional study. *Br J Dermatol* 168:762-70.

Roudsari MR, Karimi R, Mortazavian AM (2013) Health effects of probiotics on the skin. *Critical reviews in food science and nutrition*.

Saaf AM, Tengvall-Linder M, Chang HY, Adler AS, Wahlgren CF, Scheynius A, *et al.* (2008) Global expression profiling in atopic eczema reveals reciprocal expression of inflammatory and lipid genes. *PloS one* 3:e4017.

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 74:5463-7.

Schauber J, Gallo RL (2009) Antimicrobial peptides and the skin immune defense system. *J Allergy Clin Immunol* 124:R13-8.

Schittek B, Paulmann M, Senyurek I, Steffen H (2008) The role of antimicrobial peptides in human skin and in skin infectious diseases. *Infect Disord Drug Targets* 8:135-43.

Schuller E, Teichmann B, Haberstok J, Moderer M, Bieber T, Wollenberg A (2001) In situ expression of the costimulatory molecules CD80 and CD86 on langerhans cells and inflammatory dendritic epidermal cells (IDEC) in atopic dermatitis. *Archives of dermatological research* 293:448-54.

Schultz Larsen F (1993) Atopic dermatitis: a genetic-epidemiologic study in a population-based twin sample. *J Am Acad Dermatol* 28:719-23.

Schultz Larsen F, Diepgen T, Svensson A (1996) The occurrence of atopic dermatitis in north Europe: an international questionnaire study. *J Am Acad Dermatol* 34:760-4.

Schultz Larsen F, Hanifin JM (1992) Secular change in the occurrence of atopic dermatitis. *Acta dermato-venereologica Supplementum* 176:7-12.

Seguchi T, Cui CY, Kusuda S, Takahashi M, Aisu K, Tezuka T (1996) Decreased expression of filaggrin in atopic skin. *Archives of dermatological research* 288:442-6.

Seidenari S, Giusti G (1995) Objective assessment of the skin of children affected by atopic dermatitis: a study of pH, capacitance and TEWL in eczematous and clinically uninvolved skin. *Acta Derm Venereol* 75:429-33.

Sevilla LM, Nachat R, Groot KR, Klement JF, Uitto J, Djian P, *et al.* (2007) Mice deficient in involucrin, envoplakin, and periplakin have a defective epidermal barrier. *The Journal of cell biology* 179:1599-612.

Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, Kashlev M, *et al.* (2011) CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* 479:74-9.

Siegfried Z, Cedar H (1997) DNA methylation: a molecular lock. *Current biology : CB* 7:R305-7.

Silverberg JI, Hanifin JM (2013) Adult eczema prevalence and associations with asthma and other health and demographic factors: a US population-based study. *J Allergy Clin Immunol* 132:1132-8.

Skol AD, Scott LJ, Abecasis GR, Boehnke M (2006) Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nature genetics* 38:209-13.

Sly WS, Hu PY (1995) Human carbonic anhydrases and carbonic anhydrase deficiencies. *Annual Reviews Biochemistry* 64:27.

Stajminger G, Marinovic-Kulisic S, Lipozencic J, Pastar Z (2007) Most common inhalant allergens in atopic dermatitis, atopic dermatitis/allergic rhinitis, and atopic dermatitis/bronchial asthma patients: a five-year retrospective study. *Acta dermatovenerologica Croatica : ADC* 15:130-4.

Stefansson K, Brattsand M, Roosterman D, Kempkes C, Bocheva G, Steinhoff M, *et al.* (2008) Activation of proteinase-activated receptor-2 by human kallikrein-related peptidases. *The Journal of investigative dermatology* 128:18-25.

Sugarman JL, Fluhr JW, Fowler AJ, Bruckner T, Diepgen TL, Williams ML (2003) The objective severity assessment of atopic dermatitis score: an objective measure using permeability barrier function and stratum corneum hydration with computer-assisted estimates for extent of disease. *Arch Dermatol* 139:1417-22.

Sun LD, Xiao FL, Li Y, Zhou WM, Tang HY, Tang XF, *et al.* (2011) Genome-wide association study identifies two new susceptibility loci for atopic dermatitis in the Chinese Han population. *Nature genetics* 43:690-4.

Taieb A (1999) Hypothesis: from epidermal barrier dysfunction to atopic disorders. *Contact Dermatitis* 41:177-80.

Tay YK, Kong KH, Khoo L, Goh CL, Giam YC (2002) The prevalence and descriptive epidemiology of atopic dermatitis in Singapore school children. *Br J Dermatol* 146:101-6.

Tengvall K, Kierczak M, Bergvall K, Olsson M, Frankowiack M, Farias FH, *et al.* (2013) Genome-wide analysis in German shepherd dogs reveals association of a locus on CFA 27 with atopic dermatitis. *PLoS genetics* 9:e1003475.

Thomsen SF, Ulrik CS, Kyvik KO, Hjelmborg JB, Skadhauge LR, Steffensen I, *et al.* (2007) Importance of genetic factors in the etiology of atopic dermatitis: a twin study. *Allergy Asthma Proc* 28:535-9.

Tsukuba T, Okamoto K, Okamoto Y, Yanagawa M, Kohmura K, Yasuda Y, *et al.* (2003) Association of cathepsin E deficiency with development of atopic dermatitis. *Journal of biochemistry* 134:893-902.

Turul T, Tezcan I, Artac H, Bruin-Versteeg S, Barendregt B, Reisli I, *et al.* (2009) Clinical heterogeneity can hamper the diagnosis of patients with ZAP70 deficiency. *Eur J Pediatr* 168:87-93.

Valenta, R. (2000) Autoallergy: A pathogenetic factor in atopic dermatitis? *Journal of Allergy and Clinical Immunology* 105:432-7.

Valenta R, Natter S, Seiberler S, Wichlas S, Maurer D, Hess M, *et al.* (1998) Molecular characterization of an autoallergen, Hom s 1, identified by serum IgE from atopic dermatitis patients. *J Invest Dermatol* 111:1178-83.

Virtanen T, Zeiler T, Rautiainen J, Mäntyjärvi R (1999) Allergy to lipocalins: a consequence of misguided T-cell recognition of self and nonself? *Immunology Today* 20:398-400.

Voegeli R, Rawlings AV, Doppler S, Schreier T (2008) Increased basal transepidermal water loss leads to elevation of some but not all stratum corneum serine proteases. *International journal of cosmetic science* 30:435-42.

Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N (2004) Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *Journal of the National Cancer Institute* 96:434-42.

Wadonda-Kabondo N (2004) Association of parental eczema, hayfever, and asthma with atopic dermatitis in infancy: birth cohort study. *Archives of Disease in Childhood* 89:917-21.

Wang H, Li ZY, Liu Y, Persson J, Beyer I, Moller T, *et al.* (2011) Desmoglein 2 is a receptor for adenovirus serotypes 3, 7, 11 and 14. *Nature medicine* 17:96-104.

Wang WY, Barratt BJ, Clayton DG, Todd JA (2005) Genome-wide association studies: theoretical and practical concerns. *Nature reviews Genetics* 6:109-18.

Wang XS, Tan TN, Shek LP, Chng SY, Hia CP, Ong NB, *et al.* (2004) The prevalence of asthma and allergies in Singapore; data from two ISAAC surveys seven years apart. *Arch Dis Child* 89:423-6.

Weidinger S, Illig T, Baurecht H, Irvine AD, Rodriguez E, Diaz-Lacava A, *et al.* (2006) Loss-of-function variations within the filaggrin gene predispose for atopic dermatitis with allergic sensitizations. *J Allergy Clin Immunol* 118:214-9.

Weidinger S, O'Sullivan M, Illig T, Baurecht H, Depner M, Rodriguez E, *et al.* (2008) Filaggrin mutations, atopic eczema, hay fever, and asthma in children. *J Allergy Clin Immunol* 121:1203-9 e1.

Weidinger S, Willis-Owen SA, Kamatani Y, Baurecht H, Morar N, Liang L, *et al.* (2013) A genome-wide association study of atopic dermatitis identifies loci with overlapping effects on asthma and psoriasis. *Human molecular genetics* 22:4841-56.

Williams HC (2005) Clinical practice. Atopic dermatitis. *N Engl J Med* 352:2314-24.

Williams HC, Burney PG, Hay RJ, Archer CB, Shipley MJ, Hunter JJ, *et al.* (1994a) The U.K. Working Party's Diagnostic Criteria for Atopic Dermatitis. I. Derivation of a minimum set of discriminators for atopic dermatitis. *Br J Dermatol* 131:383-96.

Williams HC, Burney PG, Pembroke AC, Hay RJ (1994b) The U.K. Working Party's Diagnostic Criteria for Atopic Dermatitis. III. Independent hospital validation. *Br J Dermatol* 131:406-16.

Williams HC, Burney PG, Strachan D, Hay RJ (1994c) The U.K. Working Party's Diagnostic Criteria for Atopic Dermatitis. II. Observer variation of clinical diagnosis and signs of atopic dermatitis. *Br J Dermatol* 131:397-405.

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

Wood LC, Elias PM, Calhoun C, Tsai JC, Grunfeld C, Feingold KR (1996) Barrier disruption stimulates interleukin-1 alpha expression and release from a pre-formed pool in murine epidermis. *The Journal of investigative dermatology* 106:397-403.

Wood LC, Stalder AK, Liou A, Campbell IL, Grunfeld C, Elias PM, *et al.* (1997) Barrier disruption increases gene expression of cytokines and the 55 kD TNF receptor in murine skin. *Exp Dermatol* 6:98-104.

Wood SH, Ke X, Nuttall T, McEwan N, Ollier WE, Carter SD (2009) Genome-wide association analysis of canine atopic dermatitis and identification of disease related SNPs. *Immunogenetics* 61:765-72.

Wu Z, Hansmann B, Meyer-Hoffert U, Glaser R, Schroder JM (2009) Molecular identification and expression analysis of filaggrin-2, a member of the S100 fused-type protein family. *PloS one* 4:e5227.

Wüthrich B (2006) Natural History of Atopic Eczema. In: *Handbook of Atopic Eczema* (Ring J, Przybilla B, Ruzicka T, eds): Springer Berlin Heidelberg, 150-6.

Yamazaki M, Ishidoh K, Suga Y, Saido TC, Kawashima S, Suzuki K, *et al.* (1997) Cytoplasmic processing of human profilaggrin by active mu-calpain. *Biochemical and biophysical research communications* 235:652-6.

Yisraeli J, Frank D, Razin A, Cedar H (1988) Effect of in vitro DNA methylation on beta-globin gene expression. *Proceedings of the National Academy of Sciences of the United States of America* 85:4638-42.

Zanger P, Holzer J, Schleucher R, Steffen H, Schittek B, Gabrysch S (2009) Constitutive Expression of the Antimicrobial Peptide RNase 7 Is Associated with Staphylococcus aureus Infection of the Skin. *Journal of Infectious Diseases* 200:1907-15.

Zhou H-L, Luo G, Wise JA, Lou H (2013) Regulation of alternative splicing by local histone modifications: potential roles for RNA-guided mechanisms. *Nucleic acids research*.

Appendix: Survey questionnaire

Date: _____ Sample ID: _____

A Survey on Breathing, Nose and Skin Problems

[PLEASE ANSWER **ALL** QUESTIONS IN BLOCK LETTERS]

Personal Particulars

Full Name: _____ **Gender**: ☐M ☐F

Matriculation/Staff No.: _____ **Age**: _____ **Date of birth**: _____(dd/mm/yy)

Race: ☐Chinese ☐Malay ☐Indian ☐Other: _____(please indicate)

Country of birth: ☐Singapore ☐Other: _____ (please indicate). **No. of years in Singapore:** _____

Type of Housing: ☐HDB ☐Condominium / Private Apartments ☐Landed Property

Total monthly family income: ☐Below $2000 ☐$2000-3999 ☐$4000-5999 ☐$6000 and above

Number of people in household (including yourself): _____

Postal code: _____ **Contact No.**:_____ **Email**:_____

General Medical History

1. Weight (in Kg): _____ Height (in cm): _____

2a. Do you have any drug allergies Yes☐ No☐

IF YOU HAVE ANSWERED "NO", PLEASE SKIP TO THE NEXT QUESTION

2b Are you allergic to any of the following drugs?

Aspirin	Yes☐	No☐	
Ibuprofen	Yes☐	No☐	
Voltaren	Yes☐	No☐	
Other pain medication	Yes☐	No☐	
Other drugs	Yes☐	No☐	If Yes, please indicate:_____

3. Have you been prescribed any of the following medications by your doctor?

Antihistamines (e.g. Loratidine, Cetrizine, Chlorpheiramine)	Yes☐	No☐
Bronchodilators (e.g. Salbutamol, Terbutaline, Atrovent)	Yes☐	No☐
Inhaled steroids (e.g. Belomethasone, Budesonide, Fluticasone, Inflammide, etc)	Yes☐	No☐
Oral steroids (e.g. Prednisolone)	Yes☐	No☐
Nasal spray (e.g. Nasonex, Flixonase)	Yes☐	No☐
Topical steroids (e.g. Hydrocortisone, Betamethasone Valerate, Mometasone)	Yes☐	No☐
Antibiotic ointments (e.g. Neomycin)	Yes☐	No☐
Moisturizers / Emollients	Yes☐	No☐
Others (please indicate):_____		

4. Have you ever been told by a doctor that you have had Hepatitis A Virus (HAV) Infection? Yes☐ No☐

SECTION I

1. Have you ever had wheezing or whistling in the chest at any time in the past? Yes☐ No☐

2. Have you had wheezing or whistling in the chest in the last 12 months? Yes☐ No☐

IF YOU HAVE ANSWERED "NO" PLEASE SKIP TO QUESTION 6

3. How many attacks of wheezing have you had in the last 12 months? 4 to 12☐ More than 12☐ 1 to 3☐

4. In the past 12 months, how often, on average, has your sleep been disturbed due to wheezing?
Never woken with wheezing ☐
Less than one night per week ☐
One or more nights per week ☐

5. In the last 12 months, has wheezing ever been severe enough to limit your speech to only one or two words at a time between breaths? Yes ☐ No ☐

6. Have you ever had asthma? Yes☐ No☐

7. At what age did you first have asthma? (age) _____

8. For how many years did you have asthma? (years) _____

9. In the past 12 months, has your chest sounded wheezy during or after exercise? Yes ☐ No ☐

10. In the past 12 months, have you had a dry cough at night, apart from a cough associated with a cold or chest infection? Yes ☐ No ☐

SECTION II

*If you **had asthma in the past 12 months**, please answer the following questions.*
*If you have **"NO"** asthma, please skip this section and go to Section III*

1. In the past 12 months, how often, on average, have you experienced asthma attacks in the daytime?

Not at all ☐
Less frequently than monthly ☐
1-3 times a month ☐
1-3 times a week ☐
4-6 times a week ☐
Every day ☐

2. In the past 12 months, how often, on average, have you experienced asthma attacks at night?

Not at all ☐
Less frequently than monthly ☐
1-3 times a month ☐
1-3 times a week ☐
4-6 times a week ☐
Every day ☐

3. In the past 12 months, how many days (or part days) of school have you missed because of wheezing or asthma?

None☐
1 to 5 days ☐
6 to 10 days ☐
More than 10 days☐

4. In the past 12 months, how many times have you visited a General Practitioner's clinic or Specialist's Clinic for asthma (e.g. a wheezy episode and regular asthma checkup)?

None ☐
1 to 3 visits ☐
4 to 12 visits ☐
More than 12 visits ☐

5. In the past 12 months, how many times have you visited the Accident & Emergency Department in any hospital for asthma (e.g. a wheezy episode and regular asthma checkup)?

None ☐
1 to 3 visits ☐
4 to 12 visits ☐
More than 12 visits ☐

6. In the past 12 months, how many times have you been admitted to hospital because of wheezing or asthma?

None ☐
1 to 3 times ☐
4 to 6 times ☐
7 or more ☐

SECTION III

1. Have you ever had a problem with sneezing, or runny, or blocked nose when you **DID NOT have a cold or flu**? Yes☐ No☐

2. In the past 12 months, have you ever had a problem with sneezing or a runny or blocked nose when you **DID NOT have a cold or flu**? Yes☐ No☐

3. In the past 12 months, has **this** nose problem been accompanied by itchy-watery eyes? Yes☐ No☐

4. In the past 12 months, have you had any of the following symptoms when you **DID NOT have a cold or flu**?

PLEASE INDICATE IF YOU HAVE THE FOLLOWING SYMPTOMS (4a), HOW OFTEN YOU HAVE THEM IN A WEEK (4b), FOR HOW MANY CONSECUTIVE WEEKS YOU HAVE THE SYMPTOMS (4c), AND HOW SEVERE THE SYMPTOMS ARE (4d)

	4a		4b		4c		4d			
			≤ 3 days	≥ 4 days	≤ 3	≥ 4	Severity of symptoms*			
Symptom	Yes or No		per week	per week	weeks	weeks	0	1	2	3
Itchy nose	Yes☐	No☐	☐	☐	☐	☐	☐	☐	☐	☐
Sneezing	Yes☐	No☐	☐	☐	☐	☐	☐	☐	☐	☐
Runny nose	Yes☐	No☐	☐	☐	☐	☐	☐	☐	☐	☐
Nose blockage	Yes☐	No☐	☐	☐	☐	☐	☐	☐	☐	☐
Snore	Yes☐	No☐	☐	☐	☐	☐	☐	☐	☐	☐
Nose bleed	Yes☐	No☐	☐	☐	☐	☐	☐	☐	☐	☐

* *0 = none: no symptom evident;*
1 = mild: symptom clearly present but minimal awareness;
2 = moderate: definite awareness of symptom which is bothersome but tolerable;
3 = severe: symptom is hard to tolerate and causing interference with activities of daily life/sleeping.

5. Does your nose problem cause any following disturbances? **TICK ALL THAT APPLY**

Sleep disturbance ☐
Impairment of daily activities. Leisure and/or sport ☐
Impairment of school or work ☐
Troublesome symptoms ☐

6. How long have you been living with these nose symptoms?

Less than one year ☐
1-4 years ☐
5-10 years ☐
10 or more years ☐

7. Have you ever had allergic rhinitis? Yes☐ No ☐ Don't know☐

SECTION IV

1. Have you ever had an itchy rash which was coming and going for at least six months?	Yes☐ No☐
2. Have you had this itchy rash at any time in the last 12 months?	Yes☐ No☐
3. Has this itchy rash at any time affected any of the following places: The folds of the elbows, behind the knees, in front of the ankles, under the buttocks, or around the neck, cheeks, ears or eyes?	Yes☐ No☐

IF YOU HAVE ANSWERED "NO" PLEASE SKIP TO QUESTION 8

4. At which age did this itchy rash first occur?	Under 2 year☐ Age 2-4 years☐ Age 5 or more☐ Don't Know☐
5. Has this rash cleared completely at any time during the last 12 months?	Yes☐ No☐
6. In the past 12 months, how often, on average, have you been kept awake at night by this itchy rash?	Never in the last 12 months☐ Less than one night per week☐ One or more nights per week☐
7. In the past 12 months, have you suffered from dry skin?	Yes☐ No☐
8. Have you ever had eczema?	Yes☐ No ☐ Don't know☐

SECTION V

1. In the past 12 months, how often, on average, did you eat or drink the following?

	Never or Only occasionally	Once or Twice per week	Most or All days
Meat (e.g. Beef, lamb, chicken, pork)	☐	☐	☐
Seafood (including fish)	☐	☐	☐
Fruits	☐	☐	☐
Vegetables (green and root)	☐	☐	☐
Pulses (peas, beans, lentils)	☐	☐	☐
Cereals (including bread)	☐	☐	☐
Pasta	☐	☐	☐
Rice	☐	☐	☐
Butter	☐	☐	☐
Margarine	☐	☐	☐
Nuts	☐	☐	☐
Potatoes	☐	☐	☐
Milk	☐	☐	☐
Eggs	☐	☐	☐
Burgers / fast food	☐	☐	☐
Yakult / Vitagen / similar yogurt drinks	☐	☐	☐

2. How many times a week do you engage in vigorous physical activity long enough to make you breathe hard?

Once or twice per week☐
Most or all days☐
Never or only occasionally☐

3. How many hours of do you spend in front of the television or computer every day?

Less than 1 hour ☐
1 to 3 hours ☐
More than 3 hours to 5 hours ☐
More than 5 hours ☐

4. How often do you consume alcohol?

Frequent☐
Occasional☐
Non-drinker☐

5a. What is your smoking status?

Smoker☐
Ex-smoker☐
Non-smoker☐

IF YOU ARE A "NON-SMOKER", PLEASE SKIP TO SECTION VI
IF YOU ARE A "SMOKER", PLEASE ANSWER QUESTION 5b BELOW
IF YOU ARE AN "EX-SMOKER", PLEASE ANSWER QUESTION 5c BELOW

5b. If you currently are a smoker,
How many years have you been smoking? _____
How many cigarettes do you smoke every day? _____

5c. If you have already stopped smoking,
How many years has it been since you stopped smoking? _____
For how many years did you use to smoke? _____
How many cigarettes did you use to smoke every day? _____

SECTION VI

1. Please state the number of people living in the household who smoke cigarettes, including parents: _____

2. Does your father (or male guardian) smoke cigarettes? Yes☐ No☐

   If **YES**, please state the amount of cigerettes which your father (or male guardian) smoke each day:

   Number of Cigarettes_____ /day

3. Does your mother (or female guardian) smoke cigarettes? Yes☐ No☐

   If **YES**, please state the amount of cigarettes that your mother (or female guardian) smoke each day:

   Number of Cigarettes_____ /day

4. Does anyone living with you smoke cigarettes in your presence? Yes☐ No☐

5. Have you ever had animals (e.g. cats, dogs, mice, hamsters, etc) in your house? Yes☐ No☐

IF YOU HAVE ANSWERED "NO" PLEASE SKIP TO THE NEXT SECTION

6. At what **AGE** did you have the following animals in your house?
YOU CAN TICK MORE THAN ONE BOX

Animal \ Age	< 2	2-6	7-12	13-18	>18
Cat	☐	☐	☐	☐	☐
Dog	☐	☐	☐	☐	☐
Rodent (hamster, mouse, etc.)	☐	☐	☐	☐	☐
Birds (chicken, parrot, pigeon, etc.)	☐	☐	☐	☐	☐
Other: _____	☐	☐	☐	☐	☐

Section VII

1. What is the highest level of education completed by your mother?

Primary☐ Secondary ('O' or 'A' levels') ☐ Tertiary (Diploma, Degree or higher) ☐

2. What is the highest level of education completed by your father?

Primary☐ Secondary ('O' or 'A' levels') ☐ Tertiary (Diploma, Degree or higher) ☐

3a. Has your mother ever been diagnosed with

Asthma	Yes☐	No☐	Don't know☐
Allergic rhinitis (Hay fever)	Yes☐	No☐	Don't know☐
Eczema (Sensitive Skin)	Yes☐	No☐	Don't know☐

3b. Any drug allergies? *If any, please specify*:_____

4a. Has your father ever been diagnosed with

Asthma	Yes☐	No☐	Don't know☐
Allergic rhinitis (Hay fever)	Yes☐	No☐	Don't know☐
Eczema (Sensitive Skin)	Yes☐	No☐	Don't know☐

4b. Any drug allergies? *If any, please specify*:_____

5a. Please state the number of siblings that you have (excluding yourself): _____

5b. How many of your siblings have been diagnosed with the following?

Asthma	0☐	1☐	2☐	3 or more☐	Don't Know☐
Allergic rhinitis (Hay fever)	0☐	1☐	2☐	3 or more☐	Don't Know☐
Eczema (Sensitive skin)	0☐	1☐	2☐	3 or more☐	Don't Know☐

Thank you for taking the time to complete this questionnaire.
The information you have provided is confidential.

Name:
Matric/Staff No.:
Sample No.:

SKIN PRICK TESTING PROTOCOL

1. Prep the forearm with alcohol swab.
2. Mark 6 spots on the forearm in two parallel rows 3cm apart with a ballpoint pen.
3. Place 1.5 µl of each control and allergen extracts in the same format as listed in the table below
4. Prick the skin below the extracts with a lancet. Change the lancet after each allergen.
5. Wait for 15 minutes.
6. Measure the longest and shortest diameter of the erythema (flare) and wheal (swelling) with a ruler and if pseudopodia is present, and note it in the table below.

SKIN PRICK TEST RECORD

TEST	RESULTS (mm)	CLASS
HISTAMINE (positive control)	E x W x	
Blomia tropicalis (dust mite)	E x W x	
Dermatophagoides pteronyssinus (dust mite)	E x W x	

TEST	RESULTS (mm)	CLASS
SALINE (negative control)	E x W x	
Elaeis guineensis (oil palm pollen)	E x W x	
Curvularia spp. (fungus)	E x W x	

RESULTS **E**: Erythema flare diameter **W**: Wheal diameter

CLASS **0=** No wheal. Erythema absent or < 1mm diameter or reaction < control
1+ = Wheal absent or slight. Erythema present and < 3mm diameter
2+ = Wheal absent or slight with associated erythema > 3mm diameter
3+ = Wheal ≥ 3mm and with erythema
4+ = Wheal ≥ 3mm with pseudopodia and erythema