From THE DEPARTMENT OF LABORATORY MEDICINE
Karolinska Institutet, Stockholm, Sweden

GENETICS OF IMMUNOGLOBULIN A DEFICIENCY

CHE KANG LIM

Stockholm 2018

Cover Image: solution structure of human secretory IgA1. Sequence from PDB ID: 3CHN (Bonner, A., Almogren, A., Furtado, P.B., Kerr, M.A., Perkins, S.J. (2009) Mucosal Immunol 2: 74-84) and edited with Molsoft-ICM Software (Molsoft LLC, CA, USA)

Published by Karolinska Institutet.
Printed by Eprint AB, 2018

ISBN 978-91-7676-914-0

Genetics of Immunoglobulin A Deficiency

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Che Kang Lim

Public defense: Wednesday 10th April 2018, 10:00 AM
4U Solen, Alfred Nobels Allé 8, KI Huddinge

Principal Supervisor:
Professor Lennart Hammarström
Karolinska Institutet
Department of Laboratory Medicine
Division of Clinical Immunology & Transfusion Medicine

Co-supervisor(s):
Associate Professor Yeow Tee Goh
NUKE-NUS Medical School
Singapore General Hospital
Department of Clinical Trials & Research Centre
Department of Hematology
Division of Medicine

Opponent:
Professor Ger Rijkers
Utrecht University, University College Roosevelt
Department of Sciences

Examination Board:
Professor Ola Winqvist
Karolinska Institutet
Department of Medicine

Associate Professor Karlis Pauksens
Uppsala University
Department of Medical Sciences

Professor Anders Örn
Karolinska Institutet
Department of Microbiology, Tumor and Cell Biology (MTC)

To my dearest family

It's not that I'm so smart, it's just that I stay with problems longer

- *Albert Einstein*

人生在勤，不索何获

Sedikit-sedikit, lama-lama menjadi bukit

ABSTRACT

Immunoglobulin A deficiency (IgAD) is the most common primary immunodeficiency in Caucasian populations. It is defined as a serum IgA level below 0.07 g/L with normal IgM and IgG levels in an individual older than four years of age. Approximately one-third of these patients present with recurrent respiratory and gastrointestinal tract infections, allergic disorders and autoimmune manifestations. High familial clustering and prevalence variation by ethnicity both suggest the existence of a strong genetic component of the disease. Traditionally, IgAD has been reported as permanent, and sub-normal IgA levels remain static and persist after 20 years of observation. However, a few cases of reversion have been observed. We thus investigated the frequency of reversal in children and more than one-fifth (>20%) of Swedish children who were diagnosed before 10 years of age, reversed their IgAD status. Our observation suggests that the diagnosis of IgAD should not be made before the early teens using a cutoff level of 0.07 g/L of IgA in serum.

After suggesting improved diagnostic guidelines, we investigated the role of genetics in IgAD in a Swedish Twin cohort. Surprisingly, the prevalence of IgAD was found to be markedly increased in a twin cohort as compared with the normal Swedish adult population. Although the MHC is the main genetic factor associated with IgAD development, the MHC haplotypes were not the primary factor causing the differences observed. Nonetheless, risk-conveying MHC haplotypes including *HLA-A*01, HLA-B*08* and *HLA-DRB1*01* were found to be associated with significantly lower serum IgA concentration in the twin cohort. On the contrary, individuals who carried the protective HLA alleles *B*07, DRB1*15* and *DQB1*06* were found to have significantly higher mean IgA concentration.

We then performed a comprehensive analysis within the MHC region in order to identify the potential susceptibility genes/loci within the MHC region. In our large-scale case-control study, we identified an independent MHC haplotype (*HLA-DPB1*1301*) in the class II region associated with IgAD. In addition, MHC recombination analysis suggested a region around 110 Kbp which may contain a portion of the ancestral block. However, verification using complete sequencing did not identify any differences. Nonetheless, identification of 4310 new variants from ancestral 8.1 haplotypes will provide valuable information for the investigation of other MHC associated diseases. We also identified novel genes/variants within the MHC class III region including *AGER* (rs1800625), *RNF5* (rs3130349), *BTNL2* (rs1980493) and *HCG23* (rs3117097) that are associated with IgAD risk.

Subsequently, we investigated the association of non-MHC genes using different MHC risk haplotypes as category factors. In total, 14 different genes/loci were identified as potentially associated with IgAD in individuals carrying different MHC risk alleles, including one from *HLA-B*0801-DRB1*0301-DQB1*0201* (ancestral haplotype), three from the *HLA-DRB1*0701-DQB1*0202* cohort, two from *HLA-DRB1*01-DQB1*0501* and seven from patients who do not carry any susceptibility MHC allele. These findings suggest that the development of IgAD may be variable depending on the presence of potentially different genes within selected MHC susceptibility haplotypes that interact with the respective disease-causing non-MHC genes. Understanding the interaction between MHC and non-MHC genes and proteins may facilitate identification of the IgAD etiology.

In summary, this thesis not only helped to identify the genetic basis of IgAD, but also improved the current diagnostic definition of the disease. Further work, including protein-protein interaction investigations, gene knock-in/out and expression analyses are required to validate the functional role of the novel associations described in this thesis. As IgAD has been shown to be markedly overrepresented among patients with autoimmune diseases, further potential studies will aim to identify the link between IgAD and autoimmunity which may ultimately result in improved patient care.

LIST OF SCIENTIFIC PAPERS

I. **Reversal of Immunoglobulin A Deficiency in Children**
LIM CK, Dahle C, Elvin K, Andersson BA, Rönnelid J, Melén E, Bergström A, Truedsson L, Hammarström L.
J Clin Immunol. 2015 Jan;35(1):87-91.

II. **The higher frequency of IgA deficiency among Swedish twins is not explained by HLA haplotypes**
Frankowiack M, Kovanen RM, Repasky GA, **LIM CK**, Song C, Pedersen NL, Hammarström L.
Genes Immun 2015 Apr-May;16(3):199-205.

III. **Fine Mapping and Deep Sequencing of the Major histocompatibility complex identifies susceptibility loci/variants for Immunoglobulin A Deficiency**
Lim CK, Varadé J, Abolhassani H, Zhang T, Zhang Y, Fang M, Goh YT, Cao H, Xu X, Behrens TW, Hammarström L.
Manuscript

IV. **Distinct non-MHC gene associations of IgAD patients carrying different MHC risk alleles**
Lim CK, Varadé J, Behrens TW, Hammarström L.
Manuscript in submission

LIST OF SCIENTIFIC PAPERS NOT INCLUDED IN THIS THESIS

I. **Human TYK2 deficiency: Mycobacterial and viral infections without hyper-IgE syndrome**
Kreins AY, Ciancanelli MJ, Okada S, Kong XF, Ramírez-Alejo N, Kilic SS, El Baghdadi J, Nonoyama S, Mahdaviani SA, Ailal F, Bousfiha A, Mansouri D, Nievas E, Ma CS, Rao G, Bernasconi A, Sun Kuehn H, Niemela J, Stoddard J, Deveau P, Cobat A, El Azbaoui S, Sabri A, **Lim CK**, Sundin M, Avery DT, Halwani R, Grant AV, Boisson B, Bogunovic D, Itan Y, Moncada-Velez M, Martinez-Barricarte R, Migaud M, Deswarte C, Alsina L, Kotlarz D, Klein C, Muller-Fleckenstein I, Fleckenstein B, Cormier-Daire V, Rose-John S, Picard C, Hammarstrom L, Puel A, Al-Muhsen S, Abel L, Chaussabel D, Rosenzweig SD, Minegishi Y, Tangye SG, Bustamante J, Casanova JL, Boisson-Dupuis S.
J Exp Med. 2015 Sep 21;212(10):1641-62.

II. **Next Generation Sequencing Data Analysis in Primary Immunodeficiency Disorders - Future Directions**
Fang M, Abolhassani H, **Lim CK**, Zhang J, Hammarström L.
J Clin Immunol. 2016 May;36 Suppl 1:68-75.
.

III. **Novel genetic loci associated HLA-B*08:01 positive myasthenia gravis**
Varade J*, Wang N*, **Lim CK***, Zhang T, Zhang Y, Liu X, Piehl F, Matell R, Cao H, Xu X, Hammarström L.
J Autoimmun. 2017 Oct 13 [Epub ahead of print]

IV. **IL2RG hypomorphic mutations: Identification of a novel mutation in exon 8 and a review of the literature**
Lim CK*, Abolhassani H*, Sundin M, and Hammarström L.
Manuscript in Submission

V. **Single Center Experience in Targeted Next-Generation Sequencing for Genetic Diagnosis of Primary Immunodeficiency in China**
Xia Y, He T, Luo Y, Li C, **Lim CK**, Abolhassani H, Yang J and Hammarström L.
Manuscript in Submission

VI. **Tuberculosis and impaired IL-23-dependent IFN-γ immunity in humans homozygous for a common *TYK2* missense variant**
Boisson-Dupuis S[*], Ramirez-Alejo N[*], Zhi L[Ψ], Patin E[Ψ], Rao G[Ψ], Kerner G[Ψ], **Lim CK**[Ψ], N. Krementsov D[Ψ], Hernandez N, S. Ma C, Zhang Q, Markle J, Martinez-Barricarte R, Fish R, Deswarte C, Halpern J, Itan Y, Boisson B, Checchi A, Jabot-Hanin F, Bouaziz M, Cobat A, Pekcan S, Caliskaner Z, Inostroza J, Muller-Fleckenstein I, Fleckenstein B, Puel A, Ciancanelli M, Condino-Neto A, Strickler A, Abarca K, Teuscher C, D. Ochs H, Reisli I, H. Sayar E, El-Baghdadi J, Bustamante J, Hammarström L, G. Tangye S, Pellegrini S, Quintana-Murci L, Abel L and Casanova JL.
Manuscript in Submission

*, [Ψ] Equal contributions

CONTENTS

LIST OF ABBREVIATIONS

AA	Amino Acid
ANOVA	Analysis of variance
APC	Antigen presenting cell
BAMSE	Children, Allergy, Milieu, Stockholm, Epidemiological Survey
Bp	Base pair(s)
BWA	Burrows-Wheeler Aligner
CLEC16A	C-type lectin domain family 16A
CNV	Copy number variant
CVID	Common variable immunodeficiency
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
DZ	Dizygotic
ELISA	Enzyme-linked immunosorbent assay
GCTA-GREML	Genome-wide complex trait analysis- restricted maximum likelihood
GD	Graves' disease
GWAS	Genome-wide association study
HLA	Human leukocyte antigen
HWE	Hardy Weinberg equilibrium
IBD	Identity by descent
IFIH1	Interferon-induced helicase C domain-containing protein 1
Ig	Immunoglobulin
IgAD	Selective immunoglobulin A deficiency
IgAN	Immunoglobulin A nephropathy
InDel	Insertion/Deletion
ITP	Immune thrombocytopenic purpura
JRA	Juvenile rheumatoid arthritis
kbp	Kilobase pairs
LD	Linkage disequilibrium

Mbp	Megabase pairs
MCMC	Markov chain Monte Carlo
MG	Myasthenia gravis
MHC	Major histocompatibility complex
MS	Multiple sclerosis
MSH5	mutS homolog 5
MZ	Monozygotic
OR	Odds ratio
PBS	Phosphate buffered saline
PCR-SSP	Polymerase chain reaction-single specific primer
QC	Quality Control
RA	Rheumatoid arthritis
SBT	Sequence-based technique
SD	Standard deviation
SLE	Systemic lupus erythematosus
SNP	Single nucleotide polymorphism
T1D	Type 1 diabetes
TACI	Transmembrane activator and calcium modulator and cyclophilin ligand interactor
TNFRSF13B	Tumor necrosis factor receptor superfamily member 13B

1 INTRODUCTION

1.1 Immunoglobulin A

Immunoglobulin A (IgA) is the predominant antibody class in mucosal secretions and is the second most prevalent antibody in serum after IgG (1). IgA comprises at least 70 % of all immunoglobulins produced in the human body and plays a key role in immunity.

There are two IgA subclasses in humans, which are encoded by two functional genes, resulting in two IgA subclasses (IgA1 and IgA2, Figure 1). The length of the hinge region is the major difference between IgA1 and IgA2 (1, 2).

Whilst serum IgA predominantly consists of monomeric IgA1, secretory IgA (S-IgA) (Figure 1) is chiefly polymeric, comprising mainly dimeric forms, with an increased proportion of IgA2 (2). Serum IgA is produced by B-lymphocytes in the bone marrow while S-IgA is synthesized locally at the mucosal surfaces (2). There is a strong correlation between serum IgA concentrations and mucosal IgA concentrations (3).

The function of S-IgA is to protect the lining of the respiratory, gastrointestinal and genitourinary tracts from invading pathogens by neutralizing antigens as well as preventing the adherence of bacteria (2, 4, 5). On the other hand, the function of serum IgA is less clear, although there are reports of its involvement in triggering effector functions (2, 6).

In Swedish adults (>20 years old), the normal range of serum IgA levels for adults is 0.88-4.5 g/L (reference range at the Karolinska University Hospital Clinical Immunology/Transfusion Medicine Laboratory, Sweden).

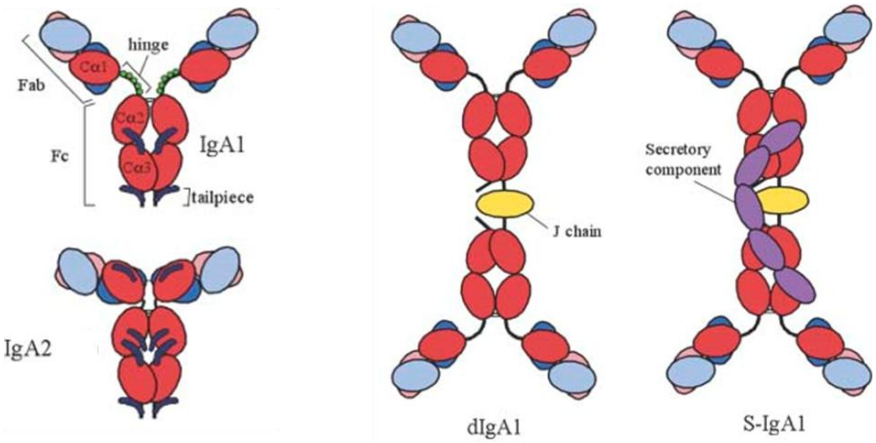

Figure 1. Monomeric forms of human IgA1, IgA2 and dimeric form of dIgA1, secretory S-IgA1. Constant regions of the heavy chain are shown in red and the variable domains of the heavy chain in pink. For the light chains, constant regions are shown in mid-blue and variable regions in pale blue. J chain is shown in yellow and secretory component in purple. On the IgA1, O-linked sugars on the IgA1 hinge are shown as green circles, while N-linked oligosaccharides are shown in dark blue. For clarity, the oligosaccharide moieties of dIgA and S-IgA1 are not shown.

This figure was adapted from Woof JM. and Kerr MA., Journal of Pathology 2006 Jan 2082: 270-82, with permission from John Wiley and Sons

1.2 IMMUNOGLOBULIN A DEFICIENCY (IGAD)

Immunoglobulin A deficiency (IgAD) is the most common human primary antibody deficiency in Caucasian populations. It is defined as a serum IgA level below or equal to 0.07 g/L in the presence of normal levels of other immunoglobulin isotypes in an individual older than four years of age (7-9).

IgAD comprises a heterogeneous group of disorders, ranging from asymptomatic (incidentally noted in healthy blood donors) to symptomatic patients presenting with different clinical phenotypes, including recurrent infections of the gastrointestinal and respiratory tracts, allergic disorders and autoimmune diseases of variable severity (10-14). In addition, recent reports have shown that patients with IgAD have significantly poorer physical health and an increased risk of early death (15, 16).

The prevalence of IgAD is variable in different populations, ranging from 1:143 in Saudi Arabia (17) to 1:18500 in Japan (18). However, the prevalence may be underestimated because some individuals with IgAD are asymptomatic, and established routine screening programs for IgAD are rare, particularly in Asia. Up to 40% of IgAD patients develop anti-IgA antibodies (19). From a transfusion medicine perspective, the presence of anti-IgA antibodies in an IgA deficient recipient is a possible cause of anaphylactic transfusion reaction. Approximately 20% of anaphylactic transfusion reactions in Western populations are associated with anti-IgA antibodies in IgA deficient recipients (20).

Since the prevalence varies widely between different populations, this supports the view that genetic factors play an essential role in the pathogenesis of the disease. In addition, the observation of familial clustering, as well as associations with known genetic loci have provided clinical evidence for the genetic predisposition to IgAD. However, the exact etiology of IgAD remains unclear.

1.3 MAJOR HISTOCOMPATIBILITY COMPLEX (MHC)

The major histocompatibility complex (MHC) locus is an extremely polymorphic region and has been one of the most intensively investigated areas in the human genome. In humans, the MHC is called the human leukocyte antigen (HLA) as the gene products were initially identified using alloantibodies against leukocytes (21). This genomic region is located on chromosome 6p21 and spans approximately 3.6 megabase pairs (Mbp) and encodes over 200 genes, many of which have a defined immunological function (22). An extended MHC of 7.6 Mb comprising more than 400 annotated genes and pseudogenes has also been described more recently (23).

The MHC complex contains two primary classes, the MHC class I and MHC class II gene clusters. The MHC class I molecules are expressed on the surface of all nucleated cells, whereas the expression of MHC class II molecules is restricted to antigen presenting cells, such as B cells and dendritic cells. The MHC class II molecules are responsible for presenting peptide antigens to T cells, thereby initiating the adaptive immune responses. In addition, there is an MHC class III region which spans the region between the MHC class I and MHC class II gene clusters. The MHC class III region contains genes that are involved in immunity, such as the complement genes and genes encoding inflammatory cytokines (24).

Given the central role of the MHC in immune function, in addition to its level of diversity, the involvement of MHC in disease susceptibility is not surprising. The importance of this locus to the pathogenesis of human disorders is proven by the reported association of polymorphisms in the HLA region with over two hundred diseases, including autoimmune diseases, primary immunodeficiency diseases, susceptibility to infections, malignancies and psychiatric conditions (25-28) (Figure 2).

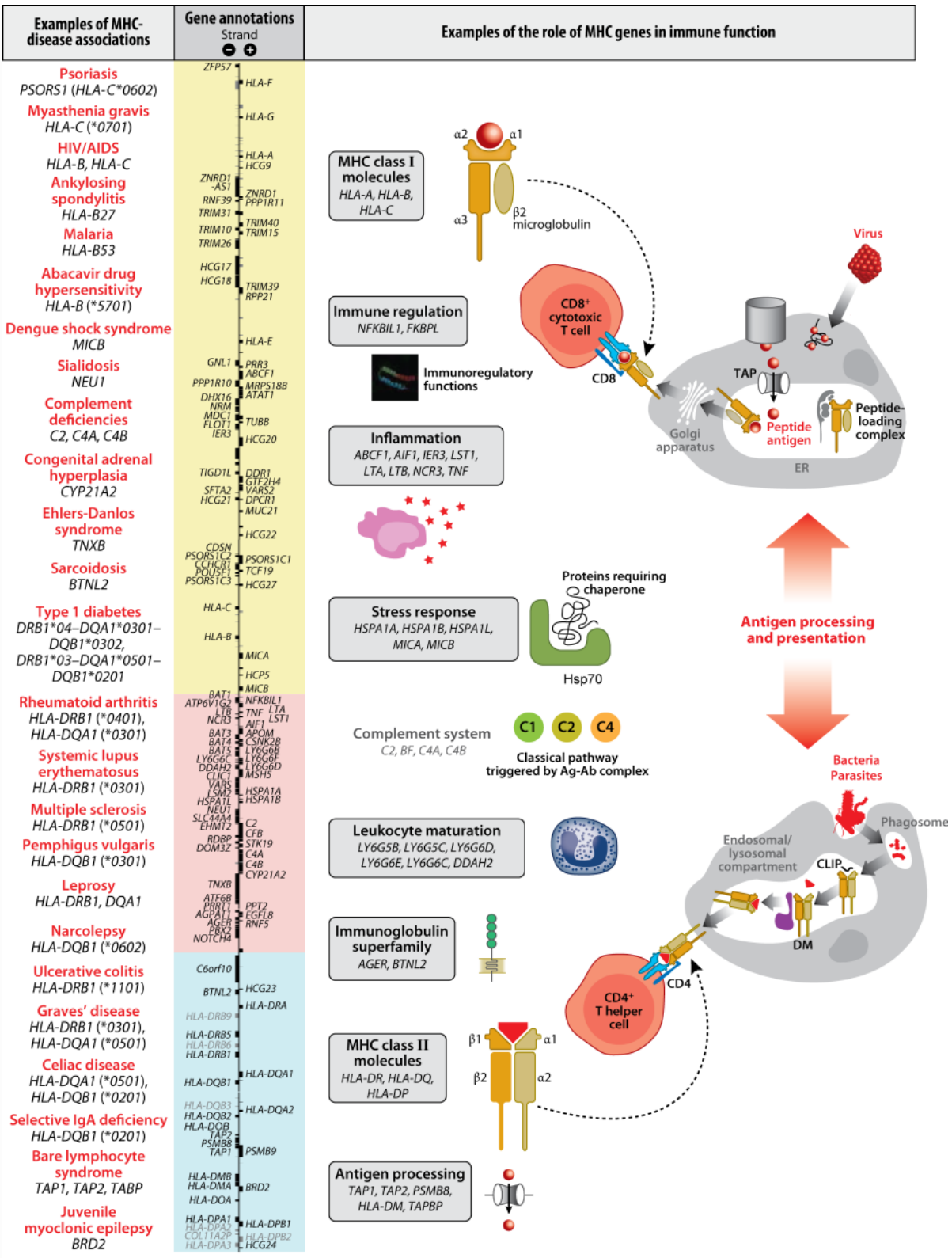

Figure 2. The MHC, disease association, and immune function. The MHC shows associations with almost all known autoimmune diseases as well as many inflammatory and infectious diseases. Major disease associations are listed by trait. Examples of the role of MHC genes in immune function are illustrated, including the key role in antigen presentation, inflammation, the complement cascade, and stress response.

This figure was adapted from Trowsdale J. and Knight JC. Annu Rev Genomics Hum Genet. 2013;14:301-23, with permission from ANNUAL REVIEWS.

1.4 MAJOR HISTOCOMPATIBILITY COMPLEX (MHC) ASSOCIATIONS OF IGAD

The MHC was first identified as a risk locus for IgAD through the association with HLA Class I and Class II markers and ascribed to specific conserved haplotypes (29-32). Most notably, the *HLA-B*08-DRB1*0301-DQB1*02* haplotype (ancestral 8.1) has been identified as the most significant genetic risk factor for IgAD in Northern European populations (33, 34). Thirteen percent of homozygous individuals with *HLA-B*0801-DRB1*0301-DQB1*0201* have been estimated to be IgA deficient (32), although this might be overestimated due to publication bias (35). Conversely, the *HLA-DRB1*1501-DQB1*06* haplotype has been shown to confer substantial protection against IgAD, with homozygous individuals showing virtually complete protection from the disease (33, 34). Additionally, susceptibility has also been shown to be associated with two other haplotypes, i.e., *HLA-B*14-DRB1*0102-DQB1*05* and the *HLA-B*44-DRB1*0701-DQB1*02* haplotypes (33, 34). Interestingly, a recent study from China showed that the ancestral 8.1 haplotype *(HLA-B*08-DRB1*0301-DQB1*02*) is also associated with the disease in the Chinese population (36, 37). Despite the strong association with the HLA locus, there has been no consensus as to the precise location of the causal variants. Some research groups have suggested that the location of the susceptibility area for IgAD is the telomeric region of MHC Class II (38-40), while others have proposed the centromeric region of MHC Class III (40-43). In addition, others have suggested a susceptibility locus in a region of MHC Class III that encodes cytokines needed for immunoglobulin production, which may also be associated with other forms of immunodeficiency, such as common variable immunodeficiency (CVID) (32). A subgroup analysis has also proposed the haplotype *MSH5-85F-DRB1*0102* at the telomeric end of MHC class II as a potential susceptibility region for IgAD (44). Furthermore, a lack of diversity of MHC Class II, and an amino acid (AA) substitution at position 57 of the *HLA-DQB1* chain may also be associated with the pathogenesis of IgAD (33).

1.5 NON-MHC ASSOCIATIONS OF IGAD

In addition to linkage to the HLA region, IgAD has been found to be associated with variants within the Interferon-induced helicase C domain-containing protein 1 gene (*IFIH1*) as well as variants within the C-type lectin domain family 16 gene (*CLEC16A*). These genes have also been implicated in the susceptibility to a variety of autoimmune disorders including type 1 diabetes (T1D), systemic lupus erythematosus (SLE) and celiac disease (CD) (45). Furthermore, recent findings show that common variants at *PVT1, ATG13-AMBRA1* and *AHI1* are associated with IgAD. These variants overlap with autoimmune markers and

correlate with 21 putative regulatory variants, such as DNase hypersensitivity sites in $FOXP3^+$ regulatory T cells (46).

1.6 CYTOGENIC DEFECTS IN IGAD

Cytogenetic defects and chromosomal abnormalities have frequently been reported in patients with IgAD. Significant anomalies were reported involving chromosomes 16 and 18 (47-52), while defects in other chromosomes have also been observed, including 4p monosomy, trisomy 10p, and translocation of 10q to 4p (53, 54). Long or short-arm deletion and ring formation have been described in some patients who exhibit additional dysmorphic features and are intellectually disabled (49). However, there are conflicting findings concerning the presence or absence of substantial chromosomal abnormalities in IgAD. A study based on asymptomatic IgAD individuals has not identified any chromosomal abnormalities (55). This observation suggests that cytogenic defects are probably not relevant to asymptomatic IgA deficient patients.

1.7 MONOGENIC MUTATIONS ASSOCIATED WITH IGAD AND ASSOCIATION WITH OTHER PRIMARY IMMUNODEFICIENCY DISEASES

The first significant monogenic mutation to be identified in IgAD was a mutation in the tumor necrosis factor receptor superfamily member 13B (*TNFRSF13B*) / transmembrane activator and calcium modulator and cyclophilin ligand interactor (TACI), which mediates isotype switching in B cells. TACI was expressed in B cells from these patients but they did not produce IgA and IgG in response to the TACI ligand, suggesting impaired isotype switching (56, 57). TACI mutations have been seen both in patients with CVID and in those with IgAD, however, are only present in a small subset of patients with each disease (56, 57). Some IgAD cases may progress to CVID (58, 59). One study suggests that individuals with a C104R, A181E or ins204A variant in the TACI gene may be at risk for disease progression from IgAD to CVID (60). Nonetheless, it is not clear whether these mutations are causative, or there may be additional defects that have not yet been identified.

Advancement of technologies such as next-generation sequencing (NGS), coupled with enhanced algorithms for bioinformatics has enabled a much broader approach to interrogate multiple genes simultaneously through one single reaction. This strategy has proven to be a practical approach by which to identify the genetic basis of Mendelian diseases in the clinical research setting. Molecular diagnostic testing using whole exome sequencing (WES) or whole genome sequencing (WGS) (61, 62) has greatly enhanced the discovery of new monogenic defects in Primary immunodeficiency diseases (PID).

Seven IgAD patients with a hypomorphic mutation in *JAK3* and *ARTEMIS (DCLRE1C)* have been reported (63, 64). In addition, Kato et al. identified a young IgAD patient with RAG1 deficiency (65) and Akhairy et al. recently identified two cases of IgAD progressing to CVID due to a mutation in the *LRBA* and *CD27* genes, respectively (66, 67).

In addition, some monogenic mutations in PID genes are associated with IgAD, including genes that are associated with combined immunodeficiency (*ATM, CHD7, DNMT3B, DKC1, MLH1, MECP2, NBS1, RAD50, PMS2, PNP, RMRP, RNF168, TINF2, TTC7, WAS* and *ZBTB24*), antibody deficiencies (*BTK, CARD11, PIK3R1, MSH2, MSH6, TACI* and *TWEAK*), defects in intrinsic and innate immunity (*CXCR4, IL12RB1* and *STAT1*) as well as genes associated with phagocytic abnormalities (*CYBB, NCF1, RAC2* and *SBDS*) (68, 69).

Furthermore, loss of function mutations in several non-MHC genes have been found in families with IgAD in which progression occurs from a normal immunologic state to IgA deficiency, with or without IgG subclass deficiency or CVID (70-73). IgAD is also reported to be associated with IgG2 subclass deficiency (74) and ataxia telangiectasia (75). Furthermore, Smith *et al*. reported a 13% (OR = 14.20) prevalence of IgAD in patients with DiGeorge syndrome (76).

1.8 RELATIONSHIP TO AUTOIMMUNITY

Various autoimmune diseases are known to be associated with IgAD, including juvenile idiopathic arthritis (JIA), dermatomyositis, sarcoidosis, Sjögren syndrome, Evans syndrome, isolated hemolytic anemia, pernicious anemia, rheumatoid arthritis (RA), SLE, Graves' disease (GD), T1D, CD, immune thrombocytopenic purpura (ITP), Hashimoto's thyroiditis, pulmonary hemosiderosis, Addison's disease, chronic nephritis, Henoch-Schonlein purpura and myasthenia gravis (MG) (77).

The same MHC haplotypes seen in patients with IgAD were found to be associated with selected autoimmune disorders, including T1D, SLE, CD, RA (12) where IgAD is markedly overrepresented (up to 30 fold). A recent study by Abolhassani *et al*. showed that autoimmune diseases were documented in approximately 30% of IgAD cases (78). This observation suggested a genetic overlap between IgAD and autoimmune disorders.

Furthermore, genome-wide association studies (GWAS) have revealed an association between IgAD and genetic variants in the gene for *IFIH1* and *CLEC16A* (45). Mutations in these genes are also associated with autoimmune diseases, which further suggest a connection between IgAD and autoimmune disease. The prevalence of autoimmune disorders has been

observed to be increased among first degree relatives of patients with IgAD, also supporting this hypothesis (79). Alternatively, the other hypothesis is that the autoimmune diseases occur as a result of recurrent infections and multiple exposures to foreign antigens that should have been "neutralized" by IgA. This hypothesis is partly demonstrated by a report which showed the presence of antibodies against milk in patients with IgA deficiency correlated with an increased frequency of serum autoantibodies (80). However, the exact pathogenetic mechanisms remain unclear.

2 AIMS

2.1 GENERAL AIM

The aim of this thesis was to study the genetic risk factors associated with IgAD in humans.

2.2 SPECIFIC AIMS

- To investigate the frequency of IgAD reversal in a large cohort of children and teenagers in order to evaluate the present definition of IgAD.

- To characterize the occurrence and concordance of IgAD in MZ and DZ twins and to identify MZ twin pairs discordant for IgAD for future functional studies.

- To fully characterize the sequence of the MHC region in order to search for mutations/variants which cause susceptibility to IgAD.

- To define etiological subgroups of IgAD using HLA risk alleles as categorical factors in order to identify potential causative non-MHC region markers involved in the pathogenesis of IgAD.

3 MATERIALS AND METHODS

3.1 SAMPLE COLLECTION

The recruitment and participation of the described Swedish patients and controls included in this thesis was approved by Regionala Etikprövningsnämnden i Stockholm (the regional ethical review board in Stockholm).

3.1.1 Children with IgAD

3.1.1.1 Symptomatic children

Between 1992 and 2012, 654 children aged from 4 to 13 years who were referred for serological testing for investigation of gastrointestinal symptoms were identified from laboratory data at five university hospitals in Sweden (283 from Karolinska Hospital in Stockholm, 78 from Sahlgrenska hospital in Gothenburg, 149 from the University Hospital in Lund, 86 from the University Hospital in Linköping and 58 from the Academic Hospital in Uppsala) and were selected for follow up analysis. Data from patients with multiple laboratory records with a minimum interval of 90 days was collected. The age of IgAD reversal was defined at the date of the first IgA value over 0.07 g/L or the first medical report indicating that IgA was detected. The follow-up concluded on 31 Jan 2014.

3.1.1.2 BAMSE follow up study

Pediatric samples from a previous study (the Children, Allergy, Milieu, Stockholm, Epidemiological survey [BAMSE] during 1994-1996) (81) were included. In total, out of 2423 children, 14 were identified to have IgAD at the age of 4 years (13) and this group had subsequently been followed up at the age of 8 years and were assessed again at 16 years of age.

3.1.2 Twin samples

For paper II, a total of 12 613 individuals (3130 MZ twins and 9483 DZ twins) were included in the study. Samples were obtained from the TwinGene project, a population-based study of Swedish twins of European ancestry born between 1911 and 1958, which included serum and DNA samples collected between 2004 and 2008 (82). For paper III and IV, 9741 Swedish twin samples (including MZ twins: 4063, one per family; DZ twins: 5678, two per family) from paper II were included, where 39 were identified as having IgAD (serum IgA < 0.07 g/L) (83).

3.1.3 Swedish IgAD patients and controls

For papers III, IV, a total of 767 Swedish IgAD patients referred to the Karolinska University Hospital in Stockholm were included in the studies. In addition, 485 healthy Swedish controls, recruited from a previous study (34) were included.

3.1.4 Multi-case family samples

For paper IV, a total of 73 Swedish multi-case families including 162 sibling pairs who had been followed up in Karolinska University Hospital in Stockholm were included in the study. Their year of birth ranged from 1910 to 2002.

3.2 SERUM IGA MEASUREMENT

3.2.1 Nephelometry

For papers I, III and IV, serum levels of IgA, IgG and IgM from the samples and controls were determined by nephelometry at the Karolinska University Hospital Clinical Immunology Laboratory, Sweden. IgAD was diagnosed if the serum IgA level was ≤ 0.07 g/L with normal serum levels of IgG and IgM.

3.2.2 Reverse-phase protein microarray

For Paper II, the serum IgA concentration was measured by a reverse-phase protein microarray as described previously (84). Briefly, diluted serum samples were spotted onto epoxy-coated microarray slides (Corning, USA) using the 2470 Arrayer Microarray Printing Platform (Aushon Biosystems, USA). Rabbit anti-human IgA (dilution 1:100 000, DakoCytomation, Denmark) and Alexa Fluor 555-conjugated goat anti-rabbit antibodies (dilution 1: 60 000, Molecular Probes, USA) were used to detect the level of IgA. Images were generated using a high-resolution microarray scanner (Agilent Technologies, USA) and subsequently analyzed by GenePix Pro 7 (Molecular Devices, USA).

3.2.3 Enzyme-linked immunosorbent assay (ELISA)

For paper I (follow up samples) and paper II (all samples), Serum IgA levels were also determined by sandwich ELISA using polyclonal rabbit anti-human IgA antibodies (DAKO, Denmark) and alkaline phosphatase-conjugated rabbit anti-human serum IgA antibodies (Jackson ImmunoResearch Laboratories, USA). Briefly, polystyrene plates were coated overnight at room temperature with 100 µl per well of the primary antibody diluted (final concentration: 1.2 mg/l) in carbonate-bicarbonate buffer (0.05 M). The plates were washed four times with phosphate-buffered saline (PBS) with 0.5% Tween20 between the incubations. All samples were titrated against a six-fold serially diluted standard, ranging

from 3.125µg/l to 100µg/l. The samples, the standard dilutions and a blank (PBS with 0.5% Tween20) were added in duplicate and incubated overnight at room temperature. The alkaline phosphatase-conjugated antibodies (final concentration: 0.3mg/l) were added and incubated for 2 hours. p-Nitrophenyl phosphate dissolved in Diethanolamine buffer (Sigma-Aldrich, USA) was used to develop the plates and the absorbance was read at 405 nm on an ELISA microplate reader (Molecular Devices, USA).

3.3 HLA TYPING

3.3.1 HLA imputation using high-resolution SNP

Imputation of HLA alleles was performed using high-resolution SNP data and two independent imputation pipelines, HLA*IMP02 (85) and SNP2HLA(86). HLA*IMP02 was carried out using a European reference panel and absolute posterior probability (Q2) ≥ 0.7 was used as the cut off for the HLA- B, HLA-DRB1 and HLA-DQB1 analyses.

For SNP2HLA, the Type 1 Diabetes Genetics Consortium dataset reference panel was used to perform the imputation of HLA types, polymorphic amino acid positions and SNPs. The minor allele frequency cut off was set at 0.5% and info score at ≥ 0.5. In total, 8404 variants, including 7125 SNPs, 1042 amino acid polymorphisms, one hundred and forty-five 4-digit resolution HLA alleles and ninety-two 2-digit resolution HLA alleles were retained for further analysis.

3.3.2 Molecular-based HLA typing

For verification of imputation results, a total of 767 Swedish IgAD samples were typed for HLA-B, HLA-DR and HLA-DQ using PCR-SSP (87) in 2-digit resolution according to the manufacturer's instructions (Olerup SSP AB, Stockholm, Sweden).

In addition, 150 IgAD samples and 25 control samples were typed at a 4-digit resolution using sequence-based typing (SBT) method as described previously (88).

3.4 EXOME SEQUENCING

Exome sequencing and the pipeline for analysis were performed as described (89). In brief, captured libraries were generated by ligation-mediated PCR. The fragments were hybridized using the Agilent SureSelect Human All Exon 50 Mb Kit (Agilent Technologies Incorporated, USA). Subsequently, each captured library was loaded onto an Illumina Hiseq2000 sequencer (Illumina Incorporated, USA) according to the manufacturer's

protocol. The high-throughput sequencing was performed to acquire the desired average sequencing depth.

3.5 GENOME-WIDE ASSOCIATION STUDY

All of the genotyping arrays were carried out using the Omni chips developed by Illumina, Inc. (Illumina Incorporated, USA) according to the manufacturer's instructions. IgAD cases were genotyped on the Omni1-Quad and Omni2.5 chips in Genentech Inc, CA, USA and the Mutation Analysis Core Facility at the Karolinska University Hospital, Stockholm, Sweden. Swedish controls were genotyped on Omni-Quad 1 (45) and the Swedish twinGene controls (82) were genotyped using the Omni Express chip. All SNPs were mapped to build hg19 coordinates using liftOver and the strand, alleles and positions were updated according to the strand data mapped to hg19 (90).

3.6 SEQUENCING OF THE MHC REGION

In order to search for new polymorphisms associated with IgAD development in the MHC region, we sequenced ten samples (5 IgAD patients and 5 controls) who carried the whole ancestral 8.1 haplotype in a homozygous form, by high-throughput sequencing, according to the protocol described previously (88). To obtain the complete MHC sequence of these samples, those regions with an average depth lower than four reads were defined as gaps and re-sequenced by Sanger sequencing. For Sanger sequencing, specific primers pairs were designed by using PRIMER3 Input version (0.4.0) software (http://frodo.wi.mit.edu/).

One hundred and fifty ng of genomic DNA was added to a final volume of 30µl of reaction mix containing the primers (0.4 mM), dNTPs (250uM), MgCl2 (2.5 mM) and 0.8 units of enzyme (GoTaq® DNA Polymerase, Promega). In order to amplify the intergenic regions, a total of 3 different PCR programs were performed according to the DNA sequence characteristics (%AT or %GC): a) Standard PCR (94ºC for 2 minute (min); 30 cycles: 94ºC for 15 second (s), 60ºC for 30 s and 72ºC for 1 min, final elongation: 72ºC for 4 min; b) Touch-Down PCR (94ºC for 5 min; 30 cycles: 94ºC for 30 s, 66ºC for 30 s and 58ºC for 12 min, final elongation: 58ºC for 30 min), the annealing temperature was decreased 0.3 ºC every cycle. c) High GC content – Touch down modified PCR (94ºC for 5 min; 40 A cycles: 94ºC for 30 s, 70ºC for 30 s and 72ºC for 40 s; 15 B cycles 94ºC for 30 s, 58ºC for 30 s and 72ºC for 40 s, final elongation: 58ºC for 5 min, the annealing temperature was decreased 0.3 ºC every A cycle, the annealing ramp was set to 33% of the speed and the elongation ramp was adjusted to 30% of the speed. The size of all fragments was confirmed by electrophoresis in 0.8% agarose gel and the specific band was sliced follow by extracted with Qiaquick gel

extraction kit (QIAGEN, Hilden, Germany) and subsequently sent for Sanger sequencing to Macrogen Inc. All the sequences were aligned to the COX ([RRID: CVCL_E534], being homozygous for *HLA-B*0801-DRB1*0301-DQB1*0201* (91)) published sequence using ClustalW software (http://www.ebi.ac.uk/Tools/msa/clustalw2) (92, 93).

3.7 BIOINFORMATICS AND STATISTICAL ANALYSIS

3.7.1 Sub-classification of population cohort and association analysis

The analysis was initiated by comparison of individuals carrying at least one MHC risk allele (*HLA*B0801-DRB1*0301-DQB1*0201* or *HLA-DRB1*0701-DQB1*0202* or *HLA-DRB1*01-DQB1*0501*) with individuals lacking a risk allele. The significantly different variants in the sample cohort were then filtered with the control cohort and only unique variants in the case comparison were considered to have an association with IgAD. The same strategy was applied in order to investigate and verify the signal by using the cohort with at least one *HLA*B0801-DRB1*0301-DQB1*0201* risk haplotype (54% of IgAD individuals). The analysis was first performed by comparing all individuals without *HLA*B0801-DRB1*0301-DQB1*0201*, followed by a comparison of individuals lacking all risk alleles.

In the subsequent analysis, cohorts homozygous for *HLA*B0801 HLA-DRB1*0301 HLA-DQB1* 0201(*68 IgAD and 123 controls), *HLA-DRB1*0701 HLA-DQB1*0202* (7 IgAD and 30 controls) and *HLA-DRB1*01 HLA-DQB1*0501* (34 IgAD and 68 controls) were selected. In addition, we also investigated cohorts homozygous for a single allele, i.e., *HLA*B0801, HLA-DRB1*0301, HLA-DRB1*0701* and *HLA-DQB1*0501*.

χ^2 tests of association for genotypes in each cohort were performed independently, using only variants that overlapped between the arrays. Variants reaching genome-wide significance ($P < 5x10^{-8}$) were considered as being significantly associated with IgAD. In addition, variants with $P < 2x10^{-7}$, and FDR< 0.05 were considered to show suggestive association with IgAD.

3.7.2 Gene-based association analysis in different subgroups

The GCTA-fastBAT software was used to analyze gene-based associations (94). A total of 24125 genes including 1522 miRNA genes from the hg19 reference were included in the analysis. MHC genes were excluded to prevent an LD effect. The gene region was set at + 50 kb from both 3' and 5' UTR of the genes. The LD cut off was set at 0.9. Non-MHC genes that had a minimum five SNPs within the regions with a total $P<2\ x10^{-6}$ were considered to be significant; while a total of $P < 2\ x10^{-4}$ were regarded as showing a suggestive association with IgAD.

3.7.3 Linkage disequilibrium proxy analysis of associated variants

LD proxy analysis was performed using LDlink software (95), and only the European population (EUR) was selected for the analysis. $R^2 > 0.9$ and $D' > 0.9$ were used as cut-offs.

3.7.4 Pre-imputation filtering

Prior to imputation, we used Genotype Harmonizer (96) to standardize the strand alignments, as well as the format for all arrays using the 1000 Genomes Project Phase 3 integrated variant set as a reference. In addition, variants with a genotyping rate <97% and evidence of deviation from Hardy-Weinberg equilibrium (HWE) in controls ($P < 1x10^{-6}$) were removed.

3.7.5 Imputation of missing genotypes in the MHC region and association analysis

MHC region imputation was performed using Impute2 (97). Only samples and controls with a high confidence HLA classification proceeded to imputation of MHC regions from chr6: 28.3 Mb to chr6: 33.8 Mb (total 180,123 SNPs) using the 1000 Genomes Project Phase 3 integrated variant set release in NCBI, with build hg19 coordinates used as the reference panel. Cases and controls were imputed together; using only genotyped variants that overlapped across all the arrays. Genotypes were imputed in 1 – 1.5 MB chunks with the effective size of the population of 20,000. Additionally, 80 haplotypes as templates when phasing observed genotypes were used. A total of 30 Markov chain Monte Carlo (MCMC) iterations were performed where the first 10 MCMC iterations were discarded as burn-in. The quality control (QC) was set as MAF > 0.5%, info score > 0.5 and calling threshold > 0.9; a total of 44857 variants passed the filtering stage and were included in the analysis. The thresholds for variant calling for each were set at 97%. In addition, the evidence of deviation from HWE ($P< 1x10^{-6}$) in controls was set. A logistic regression model was applied to test for the association and significantly associated HLA types were used as covariates in the conditional analysis (Plink (1.07 (98)). Variants reaching genome-wide significance ($P < 5x10^{-8}$) were considered as being significantly associated with IgAD.

3.7.6 Family-based association analysis

Family association analysis was performed to assess susceptibility to IgAD using a family-based association test (FBAT) (99, 100). Only HLA alleles with a minimum of ten informative families without Mendelian errors were included. Since our dataset contains multiple siblings in a family as well as various families in a pedigree, the test statistics were computed using the empirical variance, as described in Lake *et al.* (101). The single-locus analysis was conducted to test the association of IgAD with the classical HLA alleles using the additive genetic model. In addition, the multiple marker haplotype test was performed for

tightly linked HLA-DR and HLA-DQ loci. The direction and frequency of transmission were indicated by the Z statistic (Z score) and p-values.

3.7.7 Estimation of increased risk over population prevalence to siblings due to disease susceptibility HLA genes

The attributed risk analysis due to specific HLA alleles for IgAD development was based on the formula developed by Risch (102). The identity-by-descent (IBD) of all sibling pairs were calculated using the GENIBD program in the Statistical Analysis for Genetic Epidemiology (S.A.G.E) package (103). The allele frequency in the Swedish population was extracted from an allele frequency database (104). A lack of MHC recombination between HLA-DR and HLA-DQ was assumed in the calculation. The recurrence risk for siblings with IgAD was set at 50 according to a previous report (105).

3.7.8 MHC recombination analysis

The MHC region of six crossover samples, including three *HLA-DRB1*0301-HLA-DQB1*0201* samples without *HLA-B*0801* and three samples with *HLA- B*0801* only (without *DRB1*0301-HLA-DQB1*0201*) were sequenced using high-performance deep sequencing. The possible crossover layover in the class III region (the region between HLA-B and HLA-DRB1) were examined against COX, *HLA-B*0801 DRB1*0301-HLA-DQB1*0201* homozygous control and *HLA-DRB1*0301-HLA-DQB1*0201* heterozygous control using fine mapping SNP examination.

Sequencing reads of 6 crossover samples and 10 homozygous samples (including COX) were aligned to hg19 using the Burrows-Wheeler Aligner (BWA ver: 0.7.10). The bam files from 6 crossover samples were then used as input by samtools (version 1.2) phase to link heterozygous SNPs to haplotypes. Homozygous regions within these heterozygous haplotypes were also regarded as haplotypes. Ten homozygous samples were used as the control to call SNPs using Samtools software (ver1.2). The haplotypes of 6 crossover samples were filtered out if their SNPs were inconsistent with the controls, or if their length was less than 5 Kbp. Only the haplotype from the ancestral COX haplotype remained after the filtering process. Subsequently, all the samples were linked to construct the crossover regions. The final region which contained the portion of the ancestral haplotype was determined as the shared region of these areas.

3.7.9 MHC segment-based analysis

The GCTA-fastBAT analysis was performed to investigate the segment based association (94). The MHC region chr6: 28.3 – chr6: 33.8 was analyzed in 100 Kbp interval segments. In total, 242 MHC genes, including 12 miRNA genes from hg19 reference genes were included in the region. LD cut-off of 0.95 was set.

3.7.10 Heritability estimation

The genome-wide complex trait analysis-restricted maximum likelihood (GCTA-REML) software (106) was used to estimate variance explained within the MHC region using the IgAD cases and controls. The heritability of IgAD within the MHC was estimated in the Swedish IgAD cohort (n=44857 imputed genotypes) (106), using the first 10 eigenvectors as covariates in the model and adjusted for the disease prevalence in Sweden (1/600).

3.7.11 Statistical Analyses

All the statistical analyses were performed using Microsoft Excel 2010 (Microsoft, USA) and GraphPad Prism software (GraphPad, USA). Fisher exact tests and one-way analysis of variance (ANOVA) were used in paper I and paper II.

4 RESULTS

4.1 REVERSAL OF IGAD IN CHILDREN

4.1.1 Frequency of reversal of IgAD based on clinical record review

A total of 654 children diagnosed with IgAD between 1992 and 2012 at one of five participating university hospitals were identified. The children were aged between 4 and 13 years. Of these, 232 had follow-up testing results available (minimum interval 90 days) up to 31 January 2014. In order to assess the validity of using 4 years of age as the minimum age at which to reliably diagnose IgAD, the children were subdivided into three age groups (A: 4–4.99; B: 5–9.99; C: 10–12.99) based on the age at which they were first diagnosed with IgAD.

Nine out of thirty-nine (23.1%) children who were identified to have IgAD at 4 years of age had a serum IgA level above 0.07g/L upon re-testing. The average age of reversal was 9.53 ± 2.91 years. Additionally, 30 of the 131 (22.9%) children with IgAD identified between 5 and 9.99 years of age had a serum IgA level which had returned to the normal range at an average age of 12.21 ± 3.43 years of age. However, only 4 out of 62 (6.5 %) children who were diagnosed at 10 – 12.99 years of age showed a reversal of IgAD during the follow-up period. The frequency of reversal was significantly higher for children who were identified to have IgAD prior to the age of 10. No significant differences in gender were observed in IgAD reversal at any age of diagnosis.

4.1.2 BAMSE follow up study

From previous findings, five out of the ten children with IgAD had normal serum IgA levels at 8 years of age (13). Fourteen serum IgA levels from the 16-year old follow up study were analyzed using ELISA and 8 of the 14 children (57.1%) showed an increase in serum IgA above the cutoff level. Out of the eight children, four children were found to have a normalized IgA level (serum IgA over 0.7 g/L), while four remained partially deficient (serum IgA level 0.07 – 0.7 g/L), giving a prevalence of IgAD in the BAMSE cohort of 1:404 at 16 years of age.

4.2 THE HIGHER FREQUENCY OF IGA DEFICIENCY AMONG SWEDISH TWINS IS NOT EXPLAINED BY HLA HAPLOTYPES

4.2.1 Distribution of IgA concentration

The mean IgA concentration was 2.38 g/L (Standard Deviation (S.D) = 1.03) for MZ twins and 2.56 g/L (S.D=1.17) for DZ twins. The difference was significant for the mean

concentration (t = − 7.8, P= 9.9 x10^{-15}) and the effect size (η2 = 0.0047) was extremely small. The sex-stratified means were 2.53 g/L for male MZ twins and 2.26 g/L for female MZ twins, as compared with 2.71 g/L for male DZ twins and 2.43 g/L for female DZ twins. By using one-way ANOVA, the mean serum IgA concentrations for male and female twins were found to be significantly different (P= 1.43 x10^{-43}). However, no significant difference in the concentration was observed in a time-of-sampling-dependent manner.

4.2.2 Influence of HLA alleles on IgAD and IgA concentration

4.2.2.1 Influence on IgAD

Comparison of HLA haplotypes was made between the twin study population (based on imputation of HLA types from high-resolution SNP data) and data from approximately 40 000 individuals from the National Swedish Bone Marrow Donor Registry (Tobias Registret http://www.tobiasregistret.se). There was no difference in the frequency of HLA alleles in the twin cohort compared with a national population cohort. In addition, the HLA types of the MZ IgAD twins did not show significant differences when compared with the DZ twins.

4.2.2.2 IgA concentration

ANOVA was used to study the influence of individual HLA alleles on the IgA concentration. For MZ twins, IgA concentrations from only one twin from each twin pair were used, while IgA concentrations from both DZ twins were used. The IgAD associated HLA alleles *A*01, B*08, B*13, B*14, DQB1*05, DRB1*01* and *DRB1*07* were investigated. In addition, the protective alleles HLA *B*07* and *DRB1*15* (34) were also studied. From the analysis, individuals with the HLA alleles *A*01* (p =0.019), *B*08* (p=0.003), *B*14* (p=0-042), *DRB1*01* (p<0.001), *DRB1*03* (p=0.011) and *DQB1*05* (0.038) were found to have a significantly lower mean IgA concentration. On the other hand, individuals who carried the HLA alleles *B*07* (p=0.001), *DRB1*15* and *DQB1*06* (0.01) were found to have an increased mean IgA concentration. There was no significant difference observed between individuals with or without the HLA alleles *B*13, DQB1*02* and *DRB1*07*.

4.2.3 Exome sequencing of selected discordant MZ twins

Exome sequencing was performed on two selected MZ twin pairs discordant for IgAD. However, no mutations or structural variants were found in the IgA deficient twin as compared to their sibling.

4.3 FINE MAPPING AND DEEP SEQUENCING OF THE MAJOR HISTOCOMPATIBILITY COMPLEX IDENTIFIES SUSCEPTIBILITY LOCI/VARIANTS FOR IMMUNOGLOBULIN A DEFICIENCY

4.3.1 Estimation of heritability MHC effect

The genome-wide complex trait analysis-restricted maximum likelihood (GCTA-REML) software (106) was used to estimate variance explained within the MHC region using a large-scale case-control study (636 IgAD patients and 7798 controls). The heritability of IgAD in the Swedish patients in the MHC region was estimated to be 46% (standard error [S.E.] =2.9% with the first 10 eigenvectors as covariates in the model and adjusted with the disease prevalence in Sweden [0.167%]).

4.3.2 Family association analysis and estimation of attributed risk due to disease susceptibility HLA genes

Seventy-three multi-case families including 162 sibling pairs who had been followed up since 1999 were included in the analysis. Strong associations were observed in *HLA-B**08 (Z: 2.60; p=0.0094), *HLA-DR*03* (Z: 2.98; p=0.0029) and *HLA-DQ*02* (Z: 3.91; p=0.0001) while the strongest protection was seen in the *HLA-DQ6* (Z:-2.69; p=0.0072) positive individuals, which was in agreeance with a previous finding (32). We next determined the attributed risk due to this specific HLA ancestral allele for IgAD development using the identity-by-descent (IBD) of all sibling pairs based on the formula developed by Risch (102). Based on a total of 31 affected sibling pairs where at least one sibling carried the risk haplotype (*HLA- B*08-DRB1*03-DQB1*02*), the genetic effect due to this risk haplotype was 34.6%.

4.3.3 MHC fine mapping and Haplotype association analysis in a large-scale case-control study

Three class I HLA genes (*HLA-A*, *HLA-B* and *HLA-C*) and five class II HLA genes (*HLA-DRB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DPA1* and *HLA-DPB1*) were examined. Twenty-six 4-digit HLA alleles were significantly associated with IgAD ($p < 5 \times 10^{-8}$) (Figure 3). The strongest association signal was detected from *HLA-DQB1*0201* (OR = 3.35, $P = 1.13 \times 10^{-78}$) whereas the highest odds ratio (OR= 7.78, CI 95%, 5.335 – 11.34) was derived from *HLA-DRB*0102*. In addition, the multiple haplotype effect of the *HLA- B*0801-DRB1*0301-DQB1*0201* allele (OR = 3.59, $P = 3.17 \times 10^{-82}$) showed the strongest association with IgAD, while the secondary signal came from *HLA-DRB1*0701-DRB1*0202* (OR = 1.84, $P = 1.19 \times 10^{-9}$). These results were consistent with previous findings (34, 107).

Because of the extensive LD nature and complexity of the MHC region, a conditional analysis was used to identify independent MHC haplotypes that may drive IgAD risk. The most significant associated alleles from *HLA-B*, *HLA-DRB1* and *HLA-DQB1* were in the model as covariates. If any HLA alleles remained significantly associated, it would be included in the next model as a covariate until no HLA alleles would reach the stringent significant threshold ($p < 5x\ 10^{-8}$). Firstly, we used multiple logic regression analysis conditioned on the strongest risk haplotype in *HLA- B*, *HLA-DRB1* and *HLA-DQB1* alleles. The *HLA-DPB1*1301* allele (OR = 1.84, $P = 2.20 \times 10^{-9}$) showed an independent risk separate from all known associated alleles. Further conditional analysis using *HLA-DPB1*1301* as a covariate did not show any other significant associated alleles. In addition, linkage graph analysis using the Disentangler software confirmed that the *HLA-DPB1*1301* signal is independent of the currently known susceptibility haplotypes for IgAD.

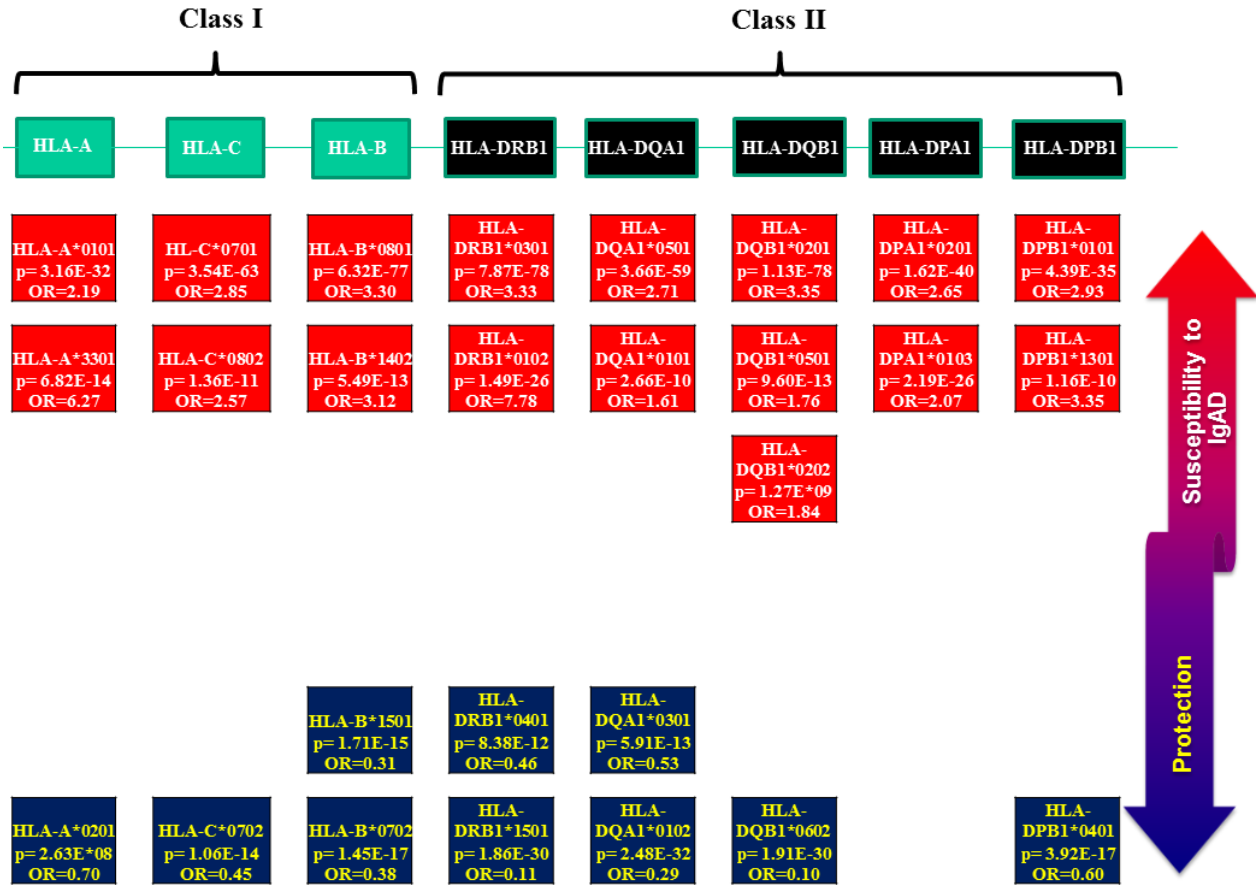

Figure 3. HLA alleles reaching significance in single locus association analysis. The positions of HLA genes show the physical order of the genes in chromosome 6 (reference: hg19).The susceptibility HLA haplotypes were highlighted in red boxes and the protection haplotypes were highlighted in blue boxes.

4.3.4 Long-range haplotype analysis

One common long-range haplotype spanning the entire HLA class I and class II regions that was in strong association with IgAD risk in the Swedish population was observed, i.e. *HLA-A*0101-C*0701-B*0801-DRB1*0301-DQA1*0501-DQB1*0201-DPA1*0201-DPB1*0101* (F case = 0.091, OR = 4.07, P= 1.71×10^{-31}). Further analysis of this long-range haplotype showed that only in the presence of *HLA-A*0101 (OR=0.488, P=* 4.56×10^{-4}*), HLA- A*0101-C*0701- B*0801(OR= 0.985,* P= 0.957) or *HLA-DRB1*0301-DQA10501-DQB1*0201 (OR= 1.03,* P= 0.913) was this haplotype not associated with IgAD.

4.3.5 MHC recombination analysis

The MHC region of crossover samples from six patients, including three patients heterozygous for *HLA-DRB1*0301-HLA-DQB1*0201* but without *HLA- B*0801*, and three samples heterozygous for *HLA-B*0801* alone but without *DRB1*0301-DQB1*0201,* where the second haplotype has been shown to be neutral with regard to IgAD were examined. The possible crossover site was located in the class III region (between *HLA-B* and *HLA-DRB1*) and was analyzed using the COX cell line (91) as a control. The haplotype blocks (5 kbp/blocks) that were homozygous in COX, but heterozygous in the samples were considered ”ancestral blocks.” A region of approximately 110 kbp in length, which may contain the shared ancestral block for all the samples was identified. The chromosome is located at chr6:31997601-32107851, within the MHC class III region. The genes located within this region are *C4B, CYP21A2, TNXB, ATF6B,* and *FKBPL*.

4.3.6 Complete sequencing of the MHC region

The MHC region was completely sequenced and aligned to the COX cell line using the ClustalW software. In total, 473 SNPs and 3837 InDels variants were identified. Thirty-nine of these were located in exons, 74 in UTR’s, 1562 in introns and 2635 in the intergenic regions which have not been listed in either the dbSNP150 database (February 03, 2017 release) or in the 1000 Genomes Project database (September 14, 2014 release) or the NHLBI-ESP project with 6500 exomes (June 07, 2013 release). There was no significant difference observed in the new variants between IgAD patients and controls.

4.3.7 MHC segment analysis and identification of the independent variants associated with IgAD

A 100K bp segment-based analysis with an LD cut-off set at 0.95 was performed. The analysis included 8434 individuals in order to investigate the strongest association region for

IgAD within the MHC region. Interestingly, the strongest association was detected in the class III region, chr6: 32.1 – 32.2, $P = 8.34 \times 10^{-110}$ whereas rs1800625 produced the strongest signal ($P_{nominal} = 4.80 \times 10^{-104}$). The stepwise conditional analysis allowed us to identify three independent signals, including *RNF5* (rs3130349), *BTNL2* (rs1980493), and *HCG23* (rs3117097) within the MHC region class III region that were associated with IgAD.

4.4 DISTINCT NON-MHC GENE ASSOCIATIONS OF IGAD PATIENTS CARRYING DIFFERENT MHC RISK ALLELES

4.4.1 MHC association analysis and multiple haplotype association investigations

8,434 samples (636 IgAD and 7798 controls) passed QC and were included in the analysis. In the MHC haplotype analysis, the *HLA-B*0801-DRB1*0301-DQB1*0201* haplotype showed the strongest association with IgAD (OR = 3.59, $P = 3.17 \times 10^{-82}$). The *HLA-DRB1*0701- DRB1*0202* (OR = 1.84, $P = 1.19 \times 10^{-9}$) and *HLA-DRB1*0101-DQB1*0501* (OR = 1.41, $P = 1.31 \times 10^{-4}$) showed a weaker association with IgAD. The effect of *HLA-DRB1*0102-DQB1*0501* was not possible to determine due to the low frequency in controls (F < 0.01). Therefore, combined signals of *HLA-DRB1*0101 DQB1*0501* and *HLA-DRB1*0102-DQB1*0501* were investigated and showed a strong association signal *HLA-DRB1*01-DQB1*0501* (OR = 1.84, $P = 3.90 \times 10^{-14}$). However, the presence of *HLA-B*0801 (*OR = 1.32, $P = 1.17 \times 10^{-1}$*)* alone or *HLA-DRB1*0301-DQB1*0201* (OR = 1.32, $P = 7.24 \times 10^{-2}$) alone was not sufficient to confer susceptibility to IgAD. Similarly, the presence of the *HLA-DRB1*0701* (OR = 1.06, $P = 7.16 \times 10^{-1}$) alone was not associated with IgAD. Since the *HLA-DQB1**0202 is in complete linkage disequilibrium (LD) with *HLA- DRB1*0701*, it was therefore not possible to investigate the effect of *HLA-DQB1*0202* without the presence of *HLA- DRB1*0701*. Similarly, the effect of *HLA-DRB1*01* or *HLA-DQB1*0501* alone could not be determined due to the low number of cases and controls.

Haplotype linkage analysis showed that *HLA-DQB1*0201* was in complete LD (100%) with *HLA-DRB1*0301* and vice versa and *HLA-DQB1*0202* had a 100% association with *HLA-DRB1*0701*. On the other hand, 74.7% of *HLA-DRB1*0701* was linked to *HLA-DQB1*0202*, whereas 24.7% was associated with *HLA-DQB1*0303*. In addition, 72% of the *HLA-DRB1*0101* and 22.8% of the *HLA-DRB1*0102* alleles were associated with *HLA-DQB1*0501*, while the remaining 5.2% were mainly associated with *HLA-DRB1*0103* and *DRB1*1001*.

4.4.2 Analysis of the influence variants outside of the MHC region in IgAD patients homozygous for high-risk HLA alleles

Based on the cross-comparison strategy, only one significantly associated non-MHC variant, rs4097492 (OR = 0.23, P= 7.63 x 10^{-9}), an intronic variant of *STXBP6* on chromosome 14 was detected in the IgAD cohort alone. Next, we analysed the individuals expressing at least one *HLA*B0801-DRB1*0301-DQB1*0201* and compared these with individuals who do not carry *HLA*B0801-DRB1*0301-DQB1*0201* (but we did include individuals carrying other risk haplotypes), and did not detect any significant variants. However, if we only compared these with individuals not carrying a risk haplotype, the identical non-MHC variants, rs4097492 (OR = 0.22, P= 4.01 x 10^{-8}) passed the genome-wide threshold (P < 5.0 x 10^{-8}).

In the *HLA-DRB1*0701-DQB1*0202* homozygous cohorts, we identified one significant (< 5x 10^{-8}) and two suggestive variants (< 2 x 10^{-7}; FDR ≤ 0.05). The peak novel variant was rs2133282 (OR = 33, P= 3.97 x 10^{-8}; FDR= 0.02), an intergenic variant located between *NOX3* and *ARID1B* on chromosome 6. Rs3917325 (OR = 59, P= 1.57 x 10^{-7}; FDR< 0.05), an UTR3 variant of *IL1R1* on chromosome 2 and rs257945 (OR = 38.67, P= 1.14 x 10^{-7}; FDR= 0.05), an intergenic variant located between *NEDD1* and *RMST* on chromosome 12 were found to be suggestively associated.

For the *HLA-DRB1*01-DQB1*0501* homozygous individuals, rs10399952 (OR= 15.4, P=5.05x10^{-9}), a variant of *FMO1* on chromosome 1 was found to be significantly associated with IgAD. However, we did not detect any strong signal (< 5x 10^{-8}or < 2 x 10^{-7}) based on the analysis method in the *HLA-B*0801-DRB1*0301-DQB1*0201* homozygous cohort.

4.4.3 Analysis of influence genes using gene-based analysis in different subgroups

In order to enhance the detection power, the data was investigated using a gene-based association analysis. With the enhanced detection method, *CD40* (P = 6.89 x 10^{-5}) on chromosome 20, with 29 SNPs in the analyzed gene region, was found to be potentially associated with IgAD in patients homozygous for *HLA-B*0801-DRB1*0301-DQB1*0201*. *DHX38* (P = 8.60 x 10^{-5}), a novel inhibitor of protein phosphatase 4 (108) located on chromosome 16 with a total of 14 SNPs in the analyzed region, was shown to be weakly associated with IgAD patients homozygous for *HLA-DRB1*01-DQB1*0501*.

A total of 7 gene regions were identified, showing suggestive association with IgAD in the patients who did not carry any of the major MHC susceptibility alleles. Most (6 out of 7) of these genes were associated with DNA repair or autoimmune disease. The associated genes

included *B3GNT6* (P = 2.09 x 10^{-6}), *TNFRSF13B* (TACI) (P = 1.12 x 10^{-4}), *GIMAP5* (P = 1.9 x 10^{-4}), *SFPQ* (P = 1.00 x 10^{-4}), *OXA1L* (P = 1.18 x 10^{-4}), *TFAP2E* (P = 6.99 x 10^{-5}), and *ZMYM4* (P = 1.07 x 10^{-4}). The locations of all the identified genes identified by two different methods are shown in the chromosome ideogram according to the MHC susceptibility groups (Figure 4).

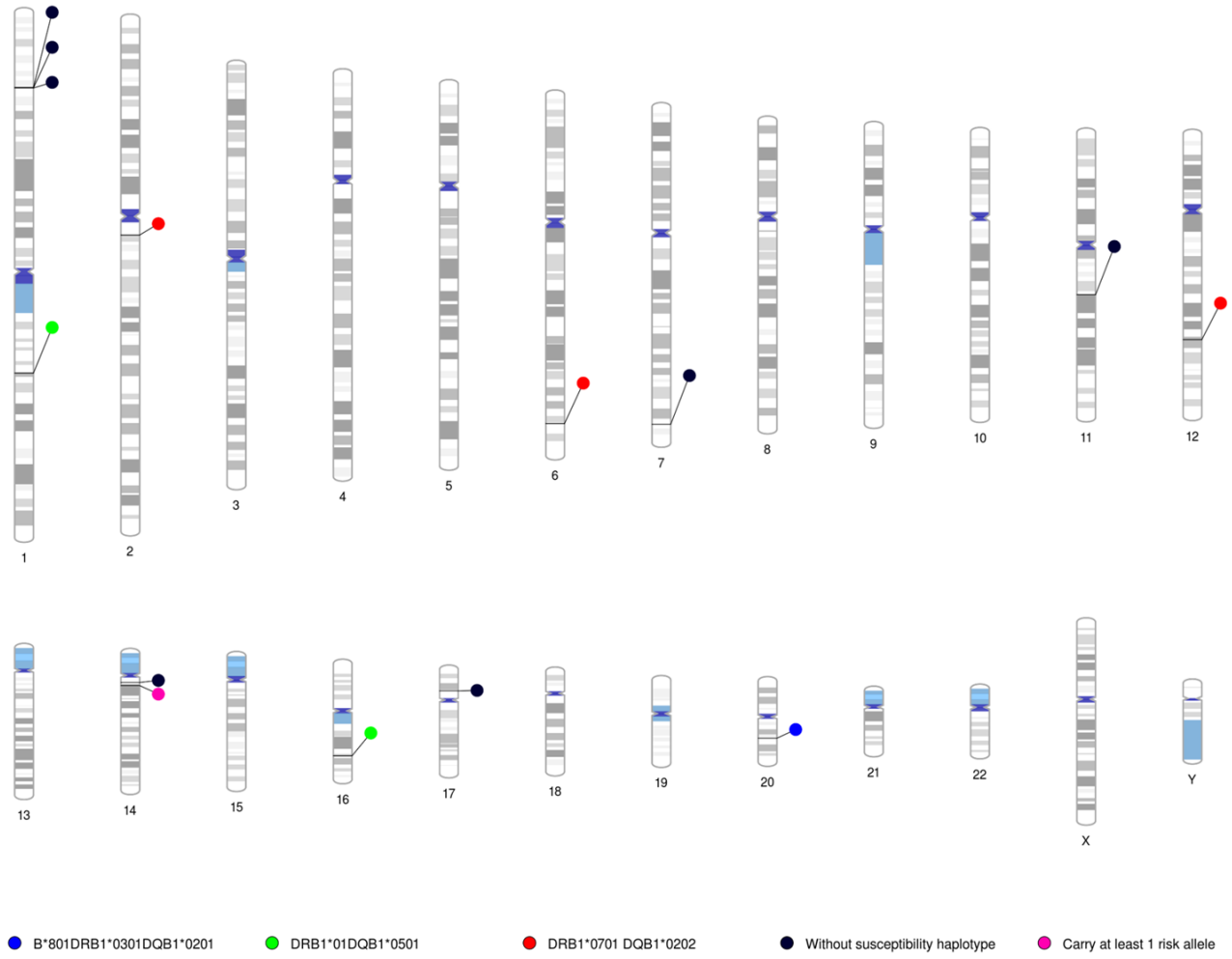

Figure 4. Chromosome ideogram for all identified susceptibility genes/locus in the patients with different MHC risk alleles. Chromosome ideogram was generated using the Phenogram software (109). Blue circle: location of genes/locus associated with the HLA-B*0801-DRB1*0301-DQB1*0201 homozygous cohort; Green circle: location of genes/locus associated with the DRB1*01-DQB1*0501 homozygous cohorts; Red circle: location of genes/locus associated with the DRB1*0701-DQB1*0202 homozygous cohorts; Black circle: location of genes/locus associated with patients carrying no MHC susceptibility genes. Pink circle: location of genes/locus associated with patients carrying at least one MHC susceptibility genes. The MHC region is highlighted in purple box. Coloured region indicated the cytogenetic band on each chromosome according to the predefined setting, based on ideogram documented in the UCSC database (110).

5 DISCUSSION AND FUTURE PERSPECTIVES

Overall, this thesis aimed to identify the genetic basis of the most common primary immunodeficiency disorder in the world. Before commencing the genetic analysis, we first verified and fine-tuned the current definition of IgAD in order to have a better sampling criteria for subsequent projects. To date, several studies have shown that compared to adults, there is a relatively high proportion of IgAD in pediatric patients that are transient in children over the age of four (72, 111-115). In **paper I**, our study was the first to evaluate reversal of IgAD among children. We showed that around 22.9% (39 out of 170) children who were diagnosed as having IgAD prior to 10 years of age had an increase in their serum IgA level with increased age. The reversal was highly significant as compared to those diagnosed after 10 years of age. This observation indicated that more than one-fifth of Swedish children might have delayed ontogeny of their IgA system. Hence, it may be too early to establish a diagnosis of IgAD using a serum level of 0.07g/L as a cutoff in children. In addition, in the BAMSE follow up study, more than 57.1 % (8 out of 14) of the children had reversed their IgAD status during a follow-up period of 12 years. The percentage of reversal was more than twice as high (57.1% vs. 22.9%) in this cohort as compared to children with a suspected disease based on gastrointestinal symptoms, celiac disease, etc. Since the BAMSE cohort represents a "healthy" control group, this observation indicated that reversal of the serum IgA level may potentially be higher in asymptomatic individuals.

Next, we initiated a twin study (**paper II**). Twins, in particular MZ twins discordant for a given disease, provide a unique opportunity to study the complex interplay of genes and environment in disease susceptibility (116). To date, there are six studies evaluating serum IgA concentrations in twins, of which three studies were case reports (117-119). Two studies quantified and compared the level of serum IgA in MZ and DZ twins (42, 43). The first paper suggested a genetic influence on serum IgA concentrations (120) whereas the second study showed that the serum IgA concentration was influenced by a combination of genetic and environmental factors (121). The last report focused on susceptibility genes for immunoglobulin deficiencies in MZ twins discordant for type 1 diabetes (122). Pairwise concordance for IgA was 50% in the MZ twins, thus markedly higher as compared with randomly paired controls.

We reported that the prevalence of IgAD was 1:241 in MZ twins and 1:198 in DZ twins, which was markedly higher than that in the healthy Swedish adult population (1:600). However, the associated HLA haplotypes were not significant in comparison with 40 000 Swedish healthy controls. This observation suggested that the genetic contribution by the

MHC region might be similar. The risk-conveying HLA alleles *A*01*, *B*08*, *B*14*, *DRB1*01* and *DQB1*05* were associated with significantly lower serum IgA concentrations in the twin cohort. Our report is the first to describe the disease risk and protective MHC alleles that influence serum IgA levels. In addition, exome sequencing from two MZ twin pairs discordant for IgAD revealed no differences between the siblings. Furthermore, the heritability of 35% for IgAD suggests a genetic influence.

We further investigated the role of genetics in IgAD by performed a comprehensive examination within the MHC region to study the IgAD risk (**paper III**) and studied the association/interaction of the non-MHC genes in individuals who carry different MHC haplotypes (**paper IV**). The majority of the twin cohort participants (77.2%) and more than 1100 IgAD patients and controls were also included in the study using fine-mapping strategies.

In general, fine-mapping strategies have confirmed the major associated locus reported by serotype analysis within a particular MHC locus in most autoimmune disorders. In addition, the strategies have provided the opportunity to investigate and determine specific allelic variants as well as independent variants in different HLA classes that are associated to the diseases. For example, for multiple sclerosis (MS), SLE and T1D, the strongest association was detected within with the MHC class II locus. However, there are also weaker associations within the class I as well as class III regions being fine-mapped (123-125). In addition, *HLA-DPB1*17*, in MHC class II, was identified being the most significant haplotype associated with dermatomyositis in Asian population (126).

By applying similar strategies, we confirmed the primary association signal from the *HLA-B*0801-DRB1*0301-DQB1*0201* haplotype (OR = 3.59, $P = 3.17 \times 10^{-82}$) and a secondary signal was observed from the *HLA-DRB1*0701- DRB1*0202* haplotype (OR = 1.84, $P = 1.19 \times 10^{-9}$). This data is in agreement with previous findings (32, 34). In addition, *HLA-DPB1*1301* (OR = 1.84, $P = 2.20 \times 10^{-9}$) was shown to be an independent risk signal and concurred with previous observations (107).

Extensive LD has been a significant obstacle to the investigation of causal genes within the MHC region. The recombination rate within the MHC is lower than the genome-wide rate determined by sperm typing (127). The rate was predicted to be less than 1 % per Mega basepair (bp) in each meiosis (128, 129). The same report showed that the recombination rate within the class II region was 0.74% and 0.94% within the class III region (128). However, only one unique case of recombination within the *DRB1-DQA1-DQB1* locus has been

observed (130). Investigation of the recombination between loci may provide clues to the location of the IgAD predisposing genes. We thus examined the MHC region in crossover samples from patients, and identified a region approximately 110 kb in length in MHC class III as shared ancestral block. However, further analysis with controls carrying the ancestral 8.1 haplotype did not show any differences. In addition, there was no significant difference noted after verification with complete sequencing of the regions (including 91 newly identified variants in the shared ancestral block). This observation suggested that the ancestral block may not contain independent IgAD predisposing genes. The MHC class I and II genes themselves may carry the susceptibility. Since not all individuals who carry MHC risk haplotypes suffer from IgAD, susceptibility to IgAD may be due to a combined effect owing to an interaction within MHC genes, as well as the interaction of an MHC-encoded susceptibility allele with a disease contributing non-MHC gene.

Fine mapping strategies have identified multiple MHC Class III variants significantly associated with autoimmune diseases, including the association to variant rs2516489 which is located between the *MICB* and *LST1* genes in MS (125). An independent association signal located at the upstream of *NOTCH4* in MHC class III region was determined by a stepwise meta-analysis of sizeable European SLE cohort (131). Additionally, an association signal of *SKIV2L* in the MHC class III region has also been suggested involving in susceptibility to the pathogenesis of SLE (132).

In addition, Goudey et al. reported 20 significant epistatic signals within the MHC which contribute to the genetic architecture of celiac disease (133). The majority of the strongest signal was located in the MHC class III region, over 1Mb upstream of the *HLA-DQA1* and *HLA-DQB1* risk loci, whereas the strongest signal corresponded to genes in the MHC class III region, in particular, *PRRC2A* and *GPANK1/C6orf47* (133). Celiac disease has been reported to be strongly associated with IgAD, with up to a 15-fold increase in the prevalence of IgAD observed among both children and adults with CD (12). The identification of strong association regions within MHC class III that likely play a role in the genetic epistasis in IgAD suggests a potentially shared genetic predisposition with celiac disease.

By using segment-based analysis in the fine-mapping strategies, the strongest association was detected in the class III region, which contains 14 genes, with rs1800625 producing the strongest signal ($P = 4.80 \times 10^{-104}$). The variant, rs1800625 is located in the promoter region of the advanced glycosylation end product-specific receptor (*AGER*). The gene is involved in innate immune mechanisms as well as mediating interactions of advanced glycosylation end

products which play a crucial role in regulating the production/expression of TNF, as well as oxidative stress (134, 135). TNF has previously been shown to be involved in the regulation of IgA production (136, 137). In addition, three independent signals including variants in *RNF5* (rs3130349), *BTNL2* (rs1980493), and *HCG23* (rs3117097) within the MHC class III region has been identified. Rs3130349 has been reported to be linked to age-related macular degeneration (AMD) (138, 139). Interestingly, in AMD patients, overactive IgA responses and increased levels of serum IgA have been observed (140). The *BTNL2* variant rs1980493 has previously been shown to be associated with autoimmune diseases including autoimmune thyroid disease, T1D and SLE (141-143) and IgAD is overrepresented among all these autoimmune disorders(12). This observation provides another piece of evidence of a potential shared genetic predisposition between IgAD and autoimmune diseases.

In addition to identifying independent variants, GWAS and MHC region fine-mapping studies allowed the analysis of epistatic interactions between genes. Multiple studies have reported the epistatic interactions between MHC and non-MHC alleles in autoimmune diseases, including T1D (144-146), GD (147-150), RA (151), SLE (146), ankylosing spondylitis (AS) (152), psoriasis (153) and MG (154). For example, the most significant epistatic interaction signal was identified between the MHC region and cytotoxic T lymphocyte antigen 4 (*CTLA4*) in the European SLE patients (155). *CTLA4* gene is upregulated in activated T cells (interacting with antigen-presenting cells (APCs)) and transmits an inhibitory signal to T cells (156). This finding indicated that proper antigen presentation and T cell activation is essential in the pathogenesis of SLE (155). In addition, the epistatic effects of MHC and non-MHC loci may also help to elucidate the mechanistic basis of the disease. For instance, individuals who carry variants in *ERAP1* showed an increased risk of psoriasis when they also carried a *HLA-C* risk allele (153). Similarly, epistatic effects were also observed for ankylosing spondylitis (AS), where loss-of-function variants of *ERAP1* reduced the risk of AS in individuals who carried the *HLA-B*27* and *HLA-B*4001* haplotypes, but not in individuals carrying other risk haplotypes (152). Animal studies have shown that *ERAP1* regulates the cleavage of relevant epitopes so that the epitopes can be presented by the *HLA-B*27* molecule (152). This observation suggested that specific epitope have to be cleaved by *ERAP1* in order to be efficiently presented. This step may be critical for identification of specific triggers for autoimmune diseases.

Our study is the first to investigate the interaction between the MHC alleles with non-MHC genes in IgAD. Altogether, we identified 14 variants/genes that are potential susceptibility loci for IgAD in different cohorts, including one for all major risk alleles, one from *HLA-*

*B*0801-DRB1*0301-DQB1*0201,* three from the *HLA-DRB1*0701-DQB1*0202* cohort, two from *HLA-DRB1*01-DQB1*0501* and seven in the patients who do not carry any susceptibility MHC alleles. The majority of the identified genes and variants were associated with an immune function or autoimmune disease, including the known IgAD associated genes, *TNFRSF13B*/*TACI* (56, 57). *TACI* gene was detected in patients who did not carry any risk alleles. On the other hand, our results show that *CD40*, which is involved in the regulation of IgA class switching (157), is suggestively associated with IgAD in the patients who are homozygous for the *HLA-B*0801-DRB1*0301-DQB1*0201* haplotype. We also identified a UTR variant in the *IL1R1* gene which was associated with the IgAD in the *HLA-DRB1*0701-DQB1*0202* cohort. Interestingly, multiple loci in *IL1R1* have been reported to be associated with IgA nephropathy (IgAN) (158), a disease which is linked to overproduction of IgA. This finding suggests that *IL1R1* may potentially play a role in IgA production.

As the clinical presentation of IgAD varies, ranging from asymptomatic to highly symptomatic patients, observation of differences in the non-MHC association in IgAD depending on the MHC risk alleles is not unexpected. The findings open up interesting perspectives for future research.

GWAS has been highly successful in the identification of genetic variants associated with complex human disorders. However, GWAS generally captures only a few percents of the estimated heritability for these complex diseases (159). Application of a conservative threshold for the association may miss some causal SNPs. In addition, it has been postulated that the missing heritability could partially be explained by epistasis or rare variants (160). A well-selected cohort may improve the discovery of missing causal signals (160). In our study, selected patients and controls enabled us to identify novel association signals within and outside the MHC regions. A similar approach has been adopted previously in other autoimmune diseases including AS and psoriasis, where the interaction between the *HLA-B*51* allele and *HLA-C* loci and *ERAP1* and *HLA-B*27,* respectively were identified (161-163). In contrast, these different strategies missed the detection of several genes which were previously implicated in the susceptibility to IgAD, including *PVT1*, *ATG13-AMBRA1*, *AHI1*, *CLEC16A* and mir-6891 (46, 164). In addition, the relatively smaller sample size in our selected cohort also limited our detection power in genes with a modest effect, although multiple strategies were employed.

In summary, this thesis has helped to advance our knowledge of the genetics of IgAD, the most common primary immunodeficiency disorder in the western world. In addition, this

study has also helped to refine the diagnostic definition of IgAD. Given the variable clinical presentation of IgAD, our findings, including 14 associated variants/genes in different cohorts outside the MHC region, four genes in the MHC class III region and three major MHC haplotypes, suggests that in contrast to other monogenic primary immunodeficiency diseases, IgAD is most likely to be a multi-complex rather than a pure monogenic disease, further evaluation of which represents the "next generation" in the genetics of immunodeficiency diseases. Further advancement in genetic studies may lead to a new definition of the disease based on the genetics of the individual. In addition, sub-categories of the disease may be identified. Currently, screening and evaluation for IgAD in the Asian population is ongoing to verify whether Caucasian haplotypes are the major risk factors for the disease. Additionally, whole genome sequencing of multi-case families is progressing in order to further enhance our understanding of the genetics of the disease. In addition, further work such as protein-protein interaction investigation, gene knock-in/out and expression testing are planned to validate the functional role of the novel associations described in the present thesis. It may also form the basis for understanding the cellular pathways involved in the pathogenesis of the disease, paving the way for potential supportive or even curative therapy for affected patients.

6 CONCLUSIONS

- In **Paper I**, the rate of reversal of IgAD in children who were diagnosed before the age of 10 is significantly high. The minimum age of four years to make a diagnosis of IgAD with a 0.07 g/L as cutoff may need to be altered to prevent a premature diagnosis of IgAD. A definitive diagnosis of IgAD should not be made before the early teens.

- In **Paper II**, the MHC haplotypes were shown not to be the factor responsible for the increased frequency of IgAD in twins. However, the MHC haplotypes may play a role in regulating serum IgA concentrations in the twin cohort. A heritability of 35% was demonstrated and suggested that genetic influences are important for IgAD.

- In **Paper III,** novel independent HLA haplotypes, as well as independent gene variants associated with IgAD disease risk within the MHC were identified. Complete sequencing of the ancestral 8.1 haplotypes identified a large number of novel genetic variants and may provide valuable information for the investigation of other MHC associated diseases.

- In **Paper IV,** the development of IgAD may be variable depending on the presence of different genes within selected MHC susceptibility haplotypes which interact with the potential disease-causing non-MHC genes. Understanding the interaction between MHC and non-MHC genes/protein may facilitate future identification of the etiology of IgAD.

7 ACKNOWLEDGEMENTS

This thesis was accomplished at the Division of Clinical Immunology and Transfusion Medicine, Department of Laboratory Medicine, Karolinska Institutet at the Karolinska University Hospital Huddinge, Sweden. It represents not only my work in the lab and at the keyboard, but it is also a milestone in more than five years of work. I would like to express my sincere gratitude to everyone who supported and guided me during the journey. In particular, with special thanks to:

Professor **Lennart Hammarström**, my principal advisor. Thank you for accepting me as a Ph.D. student in your group. I would like to express my gratitude for the encouragement, freedom, support and direction during these years. Thank you for sharing your extensive knowledge, great scientific thinking and always encouraging new ideas. Your excitement and enthusiasm for scientific research are really contagious!

Associate Professor **Yeow Tee Goh**, my co-supervisor. Thank you very much for your support, your patient, your advice and your help. It would be impossible to count all the ways that you have helped me in my career path. I will always be grateful for your guidance and kindness.

Associate Professor **Weng-Onn Lui**, my mentor. Thank you very much for useful advice, help and friendship. Thank you for the sports training and the food guide in Stockholm.

Thanks to all co-authors and collaborators for their contribution and input; without whom it would not be possible to complete this thesis.

Kerstin, appreciate your excellent support and instructions. **Renée**, salute and thanks for your fabulous support, and an exclusive guided tour in the archipelago. **Jezabel,** very grateful for your kind advice and support, I enjoy the conversation and scientific discussion with you. Thank you for proofread my thesis. **Harold,** thanks for your all the help and friendly conversation (especially "gunners") for these few years. **Hassan,** thank you, good boy, for all the help, excellent discussion and adventure together.

Jovanka, thank you for the wonderful conversations and Australia snacks, appreciate for proofreading this thesis.

Prof. **Qiang Pan-Hammarström**: Thanks for all the valuable scientific suggestions in the meeting and my seminars.

Thanks to division administrators: **Lisa, Moa, Pernilla and Naradja**, for all the fantastic job done and help me to solve all the administrative issues.

Professor **Andrej Weintraub, Arja Kramsu and Marita Ward**, thanks for your valuable advice and academic guidance.

All the Lennart's group and Pan's group members (past and present): **Rosa, Omar, Gokce, Stephan, Yin, Du, Weicheng, Chenglin, Kasper, Noel, Rozina, Radhika, Andrea, Marcel, Margarita, Apostolos, Kyriaki (Kiki), Jinqiao, Meggie, Wei Li, Xiaofei, Hui, Annica, Konstantinos, Nina, Sofia Appelberg, Mia, Zuo, Mohammad, Mahya, Georgia and Xiaogan**. It has been a pleasure working together with you; additionally, all the wonderful amenities of life, including fikas, food feast, parties and trips.

Special thanks to Assistance Professor **Charles Chuah,** thank you very much for all your support, advice and help.

All the staff in the Department of Clinical Translational Research, SGH especially Dr **Yi Zhao** for all the support and help for the projects in Singapore**;** Associate Professor **Michelle Tan** and **Ms Oi Fah Lai** for the help and advice**.** All department administrators especially **Olivia, Shahira, Jasmine,** thanks for fixing the administrative issues.

Learning center, Singapore General Hospital, very grateful to **Ms Chua Yan Qin, Ms Wee Si Min** and **Ms Agnes Chin** for the help for the past few years**.** Thank you **Singhealth Talent development fund (TDF)** for funding me.

Dr **Ang Ai Leen**, thank you for the support and help initiate the screening project in Singapore. Dr **Alvin Chua**, thanks bro for all the help and support. Very appreciate for helping me proofread the thesis. Assistance Prof **Mitchell Lai**, thanks dude for all the help and encouragement, and also the nice food guide around the world.

To the friends in Clinical Microbiology, thanks **Xiaoshan, Qian** and all the others for the wonderful conversation during lunch.

Other members of the "run more" badminton team, **Tuo, Jin, Emma, Joanna, Hao, Òscar, Corrad,** and **Pingnan**. Thanks for "running" together.

Most importantly, my family, I would like to thank my **parents** for their continuous support, unconditional love and fully confident in me. My **sisters and families**, thanks for all the joy they bring to my life. I also appreciate their help and all the supports in life.

Special gratitude to my love **Yanbing,** thank you for all your support in life. Thank you for making me feels like the luckiest man in the world every day and filling our home with love and happiness. To my little angel **Aurora,** you are the morning sunshine that keeps me awake.

Last, but not least, I would like to express my sincere gratitude to all patients and controls in our studies.

8 REFERENCES

1. Macpherson AJ, McCoy KD, Johansen FE, Brandtzaeg P. The immune geography of IgA induction and function. Mucosal immunology. 2008 Jan;1(1):11-22. PubMed PMID: 19079156.

2. Woof JM, Kerr MA. The function of immunoglobulin A in immunity. The Journal of pathology. 2006 Jan;208(2):270-82. PubMed PMID: 16362985.

3. Externest D, Meckelein B, Schmidt MA, Frey A. Correlations between antibody immune responses at different mucosal effector sites are controlled by antigen type and dosage. Infection and immunity. 2000 Jul;68(7):3830-9. PubMed PMID: 10858191. Pubmed Central PMCID: 101655.

4. Strugnell RA, Wijburg OL. The role of secretory antibodies in infection immunity. Nature reviews Microbiology. 2010 Sep;8(9):656-67. PubMed PMID: 20694027.

5. Fagarasan S, Kawamoto S, Kanagawa O, Suzuki K. Adaptive immune regulation in the gut: T cell-dependent and T cell-independent IgA synthesis. Annual review of immunology. 2010;28:243-73. PubMed PMID: 20192805.

6. Roos A, Bouwman LH, van Gijlswijk-Janssen DJ, Faber-Krol MC, Stahl GL, Daha MR. Human IgA activates the complement system via the mannan-binding lectin pathway. Journal of immunology. 2001 Sep 1;167(5):2861-8. PubMed PMID: 11509633.

7. Conley ME, Notarangelo LD, Etzioni A. Diagnostic criteria for primary immunodeficiencies. Representing PAGID (Pan-American Group for Immunodeficiency) and ESID (European Society for Immunodeficiencies). Clinical immunology. 1999 Dec;93(3):190-7. PubMed PMID: 10600329.

8. Al-Herz W, Bousfiha A, Casanova JL, Chapel H, Conley ME, Cunningham-Rundles C, et al. Primary immunodeficiency diseases: an update on the classification from the international union of immunological societies expert committee for primary immunodeficiency. Frontiers in immunology. 2011;2:54. PubMed PMID: 22566844. Pubmed Central PMCID: 3342372.

9. Picard C, Al-Herz W, Bousfiha A, Casanova JL, Chatila T, Conley ME, et al. Primary Immunodeficiency Diseases: an Update on the Classification from the International Union of Immunological Societies Expert Committee for Primary Immunodeficiency 2015. Journal of clinical immunology. 2015 Nov;35(8):696-726. PubMed PMID: 26482257. Pubmed Central PMCID: 4659841.

10. Aghamohammadi A, Cheraghi T, Gharagozlou M, Movahedi M, Rezaei N, Yeganeh M, et al. IgA deficiency: correlation between clinical and immunological phenotypes. Journal of clinical immunology. 2009 Jan;29(1):130-6. PubMed PMID: 18683032.

11. Jorgensen GH, Gardulf A, Sigurdsson MI, Sigurdardottir ST, Thorsteinsdottir I, Gudmundsson S, et al. Clinical symptoms in adults with selective IgA deficiency: a case-control study. Journal of clinical immunology. 2013 May;33(4):742-7. PubMed PMID: 23389234.

12. Wang N, Shen N, Vyse TJ, Anand V, Gunnarson I, Sturfelt G, et al. Selective IgA deficiency in autoimmune diseases. Molecular medicine. 2011;17(11-12):1383-96. PubMed PMID: 21826374. Pubmed Central PMCID: 3321806.

13. Janzi M, Kull I, Sjoberg R, Wan J, Melen E, Bayat N, et al. Selective IgA deficiency in early life: association to infections and allergic diseases during childhood. Clinical immunology. 2009 Oct;133(1):78-85. PubMed PMID: 19541543.

14. Vo Ngoc DT, Krist L, van Overveld FJ, Rijkers GT. The long and winding road to IgA deficiency: causes and consequences. Expert review of clinical immunology. 2017 Apr;13(4):371-82. PubMed PMID: 27776452.

15. Ludvigsson JF, Neovius M, Hammarstrom L. IgA deficiency and mortality: a population-based cohort study. Journal of clinical immunology. 2013 Nov;33(8):1317-24. PubMed PMID: 24122027.

16. Jorgensen GH, Gardulf A, Sigurdsson MI, Arnlaugsson S, Hammarstrom L, Ludviksson BR. Health-related quality of life (HRQL) in immunodeficient adults with selective IgA deficiency compared with age- and gender-matched controls and identification of risk factors for poor HRQL. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2014 Mar;23(2):645-58. PubMed PMID: 24022790.

17. al-Attas RA, Rahi AH. Primary antibody deficiency in Arabs: first report from eastern Saudi Arabia. Journal of clinical immunology. 1998 Sep;18(5):368-71. PubMed PMID: 9793829.

18. Kanoh T, Mizumoto T, Yasuda N, Koya M, Ohno Y, Uchino H, et al. Selective IgA deficiency in Japanese blood donors: frequency and statistical analysis. Vox sanguinis. 1986;50(2):81-6. PubMed PMID: 3485858.

19. Hammarstrom L, Persson MA, Smith CI. Anti-IgA in selective IgA deficiency. In vitro effects and Ig subclass pattern of human anti-IgA. Scandinavian journal of immunology. 1983 Dec;18(6):509-13. PubMed PMID: 6607511.

20. Vassallo RR. Review: IgA anaphylactic transfusion reactions. Part I. Laboratory diagnosis, incidence, and supply of IgA-deficient products. Immunohematology / American Red Cross. 2004;20(4):226-33. PubMed PMID: 15679454.

21. Thorsby E. A short history of HLA. Tissue antigens. 2009 Aug;74(2):101-16. PubMed PMID: 19523022.

22. The MHC sequencing consortium. Complete sequence and gene map of a human major histocompatibility complex. Nature. 1999 Oct 28;401(6756):921-3. PubMed PMID: 10553908.

23. Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK, et al. Gene map of the extended human MHC. Nature reviews Genetics. 2004 Dec;5(12):889-99. PubMed PMID: 15573121.

24. Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. Nature. 1999 Oct 28;401(6756):921-3. PubMed PMID: 10553908.

25. Klein J, Sato A. The HLA System- Second of Two Parts. The New England journal of medicine. 2000 September 14, 2000;343(11):782-6.

26. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic acids research. 2014 Jan;42(Database issue):D1001-6. PubMed PMID: 24316577. Pubmed Central PMCID: 3965119.

27. Matzaraki V, Kumar V, Wijmenga C, Zhernakova A. The MHC locus and genetic susceptibility to autoimmune and infectious diseases. Genome biology. 2017 Apr 27;18(1):76. PubMed PMID: 28449694. Pubmed Central PMCID: 5406920.

28. Trowsdale J, Knight JC. Major histocompatibility complex genomics and human disease. Annual review of genomics and human genetics. 2013;14:301-23. PubMed PMID: 23875801. Pubmed Central PMCID: 4426292.

29. Ambrus M, Hernadi E, Bajtai G. Prevalence of HLA-A1 and HLA-B8 antigens in selective IgA deficiency. Clin Immunol Immunopathol. 1977 May;7(3):311-4. PubMed PMID: 872455. eng.

30. Cobain TJ, French MA, Christiansen FT, Dawkins RL. Association of IgA deficiency with HLA A28 and B14. Tissue Antigens. 1983 Aug;22(2):151-4. PubMed PMID: 6623450. eng.

31. Hammarstrom L, Smith CI. HLA-A, B, C and DR antigens in immunoglobulin A deficiency. Tissue Antigens. 1983 Jan;21(1):75-9. PubMed PMID: 6601317.

32. Schroeder HW, Jr., Zhu ZB, March RE, Campbell RD, Berney SM, Nedospasov SA, et al. Susceptibility locus for IgA deficiency and common variable immunodeficiency in the HLA-DR3, -B8, -A1 haplotypes. Molecular medicine. 1998 Feb;4(2):72-86. PubMed PMID: 9508785.

33. Olerup O, Smith CI, Hammarstrom L. Different amino acids at position 57 of the HLA-DQ beta chain associated with susceptibility and resistance to IgA deficiency. Nature. 1990 Sep 20;347(6290):289-90. PubMed PMID: 1976229.

34. Ferreira RC, Pan-Hammarstrom Q, Graham RR, Fontan G, Lee AT, Ortmann W, et al. High-density SNP mapping of the HLA region identifies multiple independent susceptibility loci associated with selective IgA deficiency. PLoS genetics. 2012 Jan;8(1):e1002476. PubMed PMID: 22291608. Pubmed Central PMCID: 3266887.

35. Mohammadi J, Ramanujam R, Jarefors S, Rezaei N, Aghamohammadi A, Gregersen PK, et al. IgA Deficiency and the MHC: assessment of relative risk and microheterogeneity within the HLA A1 B8, DR3 (8.1) haplotype. Journal of clinical immunology. 2009 Oct 16. PubMed PMID: 19834793. Eng.

36. Wang N, Lu P, Ling B, Zhu Z, Hammarstrom L. Caucasian origin of disease associated HLA haplotypes in chinese blood donors with IgA deficiency. Journal of clinical immunology. 2014 Feb;34(2):157-62. PubMed PMID: 24402621.

37. Lu P, Ling B, Wang N, Hammarstrom L. [Study on immunoglobulin A Deficiency(IgAD) in Chinese Shanghai Blood Donors]. Zhongguo shi yan xue ye xue za zhi. 2016 Aug;24(4):1216-20. PubMed PMID: 27531803.

38. Olerup O, Smith CI, Bjorkander J, Hammarstrom L. Shared HLA class II-associated genetic susceptibility and resistance, related to the HLA-DQB1 gene, in IgA deficiency and common variable immunodeficiency. Proc Natl Acad Sci U S A. 1992 Nov 15;89(22):10653-7. PubMed PMID: 1438261. eng.

39. Kralovicova J, Hammarstrom L, Plebani A, Webster AD, Vorechovsky I. Fine-scale mapping at IGAD1 and genome-wide genetic linkage analysis implicate HLA-DQ/DR as a major susceptibility locus in selective IgA deficiency and common variable immunodeficiency. J Immunol. 2003 Mar 1;170(5):2765-75. PubMed PMID: 12594308.

40. Vorechovsky I, Cullen M, Carrington M, Hammarstrom L, Webster AD. Fine mapping of IGAD1 in IgA deficiency and common variable immunodeficiency:

identification and characterization of haplotypes shared by affected members of 101 multiple-case families. Journal of immunology. 2000 Apr 15;164(8):4408-16. PubMed PMID: 10754342.

41. Schaffer FM, Palermos J, Zhu ZB, Barger BO, Cooper MD, Volanakis JE. Individuals with IgA deficiency and common variable immunodeficiency share polymorphisms of major histocompatibility complex class III genes. Proc Natl Acad Sci U S A. 1989 Oct;86(20):8015-9. PubMed PMID: 2573059.

42. Volanakis JE, Zhu ZB, Schaffer FM, Macon KJ, Palermos J, Barger BO, et al. Major histocompatibility complex class III genes and susceptibility to immunoglobulin A deficiency and common variable immunodeficiency. J Clin Invest. 1992 Jun;89(6):1914-22. PubMed PMID: 1351062. eng.

43. Cucca F, Zhu ZB, Khanna A, Cossu F, Congia M, Badiali M, et al. Evaluation of IgA deficiency in Sardinians indicates a susceptibility gene is encoded within the HLA class III region. Clin Exp Immunol. 1998 Jan;111(1):76-80. PubMed PMID: 9472664.

44. Pozo ND, Medrano LM, Cenit MC, Fernandez-Arquero M, Ferreira A, Garcia-Rodriguez MC, et al. MSH5 is not a genetic predisposing factor for immunoglobulin A deficiency but marks the HLA-DRB1*0102 subgroup carrying susceptibility. Human immunology. 2010 Sep;71(9):861-4. PubMed PMID: 20542071.

45. Ferreira RC, Pan-Hammarstrom Q, Graham RR, Gateva V, Fontan G, Lee AT, et al. Association of IFIH1 and other autoimmunity risk alleles with selective IgA deficiency. Nature genetics. 2010 Sep;42(9):777-80. PubMed PMID: 20694011.

46. Bronson PG, Chang D, Bhangale T, Seldin MF, Ortmann W, Ferreira RC, et al. Common variants at PVT1, ATG13-AMBRA1, AHI1 and CLEC16A are associated with selective IgA deficiency. Nature genetics. 2016 Nov;48(11):1425-9. PubMed PMID: 27723758.

47. Gilgenkrantz S, Charles JM, Cabrol C, Mauuary G, Vigneron C. [Deletion of the short arm of chromosome 18 due to t(22-;18p+) translocation with IgA deficiency. Cytogenetic study with autoradiography and fluorescence]. Annales de genetique. 1972 Dec;15(4):275-81. PubMed PMID: 4539488. Deletion du bras court du 18 par translocation t(22-;18p+) avec deficit en IgA. Etude cytogenetique avec autoradiographie et fluorescence.

48. Murken JD, Salzer G, Kunze D. [Ring-chromosome 18 and IgA deficiency in a 6-year-old girl (46,XX,18r)]. Zeitschrift fur Kinderheilkunde. 1970;109(1):1-10. PubMed PMID: 5478808. Ringchromosom Nr. 18 und fehlendes IgA bei einem 6jahrigen Madchen (46,XX,18r).

49. Ogata K, Iinuma K, Kammura K, Morinaga R, Kato J. A case report of a presumptive +i(18p) associated with serum IgA deficiency. Clinical genetics. 1977 Mar;11(3):184-8. PubMed PMID: 837568.

50. Schaffer AA, Pfannstiel J, Webster AD, Plebani A, Hammarstrom L, Grimbacher B. Analysis of families with common variable immunodeficiency (CVID) and IgA deficiency suggests linkage of CVID to chromosome 16q. Human genetics. 2006 Feb;118(6):725-9. PubMed PMID: 16328471. Pubmed Central PMCID: 1385708.

51. Lewkonia RM, Lin CC, Haslam RH. Selective IgA deficiency with 18q+ and 18q-- karyotypic anomalies. Journal of medical genetics. 1980 Dec;17(6):453-6. PubMed PMID: 6782212. Pubmed Central PMCID: 1885921.

52. Dostal A, Linnankivi T, Somer M, Kahkonen M, Litzman J, Tienari P. Mapping susceptibility gene locus for IgA deficiency at del(18)(q22.3-q23); report of familial cryptic chromosome t(18q; 10p) translocations. International journal of immunogenetics. 2007 Jun;34(3):143-7. PubMed PMID: 17504501.

53. Saiga T, Hashimoto K, Kimura N, Ono H, Hiai H. Trisomy 10p and translocation of 10q to 4p associated with selective dysgenesis of IgA-producing cells in lymphoid tissue. Pathology international. 2007 Jan;57(1):37-42. PubMed PMID: 17199741.

54. Hanley-Lopez J, Estabrooks LL, Stiehm R. Antibody deficiency in Wolf-Hirschhorn syndrome. The Journal of pediatrics. 1998 Jul;133(1):141-3. PubMed PMID: 9672528.

55. Herrmann RP, Chipper L, Bell S. Chromosomal studies in healthy blood donors with IgA deficiency. Clinical genetics. 1982 Nov;22(5):231-3. PubMed PMID: 7151309.

56. Castigli E, Wilson SA, Garibyan L, Rachid R, Bonilla F, Schneider L, et al. TACI is mutant in common variable immunodeficiency and IgA deficiency. Nature genetics. 2005 Aug;37(8):829-34. PubMed PMID: 16007086.

57. Rachid R, Castigli E, Geha RS, Bonilla FA. TACI mutation in common variable immunodeficiency and IgA deficiency. Current allergy and asthma reports. 2006 Sep;6(5):357-62. PubMed PMID: 16899196.

58. Espanol T, Catala M, Hernandez M, Caragol I, Bertran JM. Development of a common variable immunodeficiency in IgA-deficient patients. Clinical immunology and immunopathology. 1996 Sep;80(3 Pt 1):333-5. PubMed PMID: 8811056.

59. Aghamohammadi A, Mohammadi J, Parvaneh N, Rezaei N, Moin M, Espanol T, et al. Progression of selective IgA deficiency to common variable immunodeficiency. International archives of allergy and immunology. 2008;147(2):87-92. PubMed PMID: 18520152.

60. Pan-Hammarstrom Q, Salzer U, Du L, Bjorkander J, Cunningham-Rundles C, Nelson DL, et al. Reexamining the role of TACI coding variants in common variable immunodeficiency and selective IgA deficiency. Nature genetics. 2007 Apr;39(4):429-30. PubMed PMID: 17392797. Pubmed Central PMCID: 2931279.

61. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The next-generation sequencing revolution and its impact on genomics. Cell. 2013 Sep 26;155(1):27-38. PubMed PMID: 24074859. Pubmed Central PMCID: 3969849.

62. Nijman IJ, van Montfrans JM, Hoogstraat M, Boes ML, van de Corput L, Renner ED, et al. Targeted next-generation sequencing: a novel diagnostic tool for primary immunodeficiencies. The Journal of allergy and clinical immunology. 2014 Feb;133(2):529-34. PubMed PMID: 24139496.

63. Roberts JL, Lengi A, Brown SM, Chen M, Zhou YJ, O'Shea JJ, et al. Janus kinase 3 (JAK3) deficiency: clinical, immunologic, and molecular analyses of 10 patients and outcomes of stem cell transplantation. Blood. 2004 Mar 15;103(6):2009-18. PubMed PMID: 14615376.

64. Volk T, Pannicke U, Reisli I, Bulashevska A, Ritter J, Bjorkman A, et al. DCLRE1C (ARTEMIS) mutations causing phenotypes ranging from atypical severe combined immunodeficiency to mere antibody deficiency. Human molecular genetics. 2015 Dec 20;24(25):7361-72. PubMed PMID: 26476407. Pubmed Central PMCID: 4664172.

65. Kato T, Crestani E, Kamae C, Honma K, Yokosuka T, Ikegawa T, et al. RAG1 deficiency may present clinically as selective IgA deficiency. Journal of clinical immunology. 2015 Apr;35(3):280-8. PubMed PMID: 25739914.

66. Alkhairy OK, Perez-Becker R, Driessen GJ, Abolhassani H, van Montfrans J, Borte S, et al. Novel mutations in TNFRSF7/CD27: Clinical, immunologic, and genetic characterization of human CD27 deficiency. The Journal of allergy and clinical immunology. 2015 Sep;136(3):703-12 e10. PubMed PMID: 25843314.

67. Alkhairy OK, Abolhassani H, Rezaei N, Fang M, Andersen KK, Chavoshzadeh Z, et al. Spectrum of Phenotypes Associated with Mutations in LRBA. Journal of clinical immunology. 2016 Jan;36(1):33-45. PubMed PMID: 26707784.

68. Abolhassani H, Aghamohammadi A, Hammarstrom L. Monogenic mutations associated with IgA deficiency. Expert review of clinical immunology. 2016 Jun 21:1-15. PubMed PMID: 27266541.

69. Yazdani R, Azizi G, Abolhassani H, Aghamohammadi A. Selective IgA deficiency: Epidemiology, Pathogenesis, Clinical Phenotype, Diagnosis, Prognosis and Management. Scandinavian journal of immunology. 2016 Oct 20. PubMed PMID: 27763681.

70. Koistinen J. Familial clustering of selective IgA deficiency. Vox sanguinis. 1976;30(3):181-90. PubMed PMID: 1251582.

71. Vassallo CL, Zawadzki ZA, Simons JR. Recurrent respiratory infections in a family with immunoglobulin A deficiency. The American review of respiratory disease. 1970 Feb;101(2):245-51. PubMed PMID: 4188959.

72. Johnson ML, Keeton LG, Zhu ZB, Volanakis JE, Cooper MD, Schroeder HW, Jr. Age-related changes in serum immunoglobulins in patients with familial IgA deficiency and common variable immunodeficiency (CVID). Clinical and experimental immunology. 1997 Jun;108(3):477-83. PubMed PMID: 9182895. Pubmed Central PMCID: 1904696.

73. Vorechovsky I, Webster AD, Plebani A, Hammarstrom L. Genetic linkage of IgA deficiency to the major histocompatibility complex: evidence for allele segregation distortion, parent-of-origin penetrance differences, and the role of anti-IgA antibodies in disease predisposition. American journal of human genetics. 1999 Apr;64(4):1096-109. PubMed PMID: 10090895. Pubmed Central PMCID: 1377834.

74. Sandler SG, Trimble J, Mallory DM. Coexistent IgG2 and IgA deficiencies in blood donors. Transfusion. 1996 Mar;36(3):256-8. PubMed PMID: 8604512.

75. Birrell GW, Kneebone K, Nefedov M, Nefedova E, Jartsev MN, Mitsui M, et al. ATM mutations, haplotype analysis, and immunological status of Russian patients with ataxia telangiectasia. Human mutation. 2005 Jun;25(6):593. PubMed PMID: 15880721.

76. Smith CA, Driscoll DA, Emanuel BS, McDonald-McGinn DM, Zackai EH, Sullivan KE. Increased prevalence of immunoglobulin A deficiency in patients with the chromosome 22q11.2 deletion syndrome (DiGeorge syndrome/velocardiofacial syndrome). Clinical and diagnostic laboratory immunology. 1998 May;5(3):415-7. PubMed PMID: 9606003. Pubmed Central PMCID: 104536.

77. Robert R. Rich TAF, William T. Shearer, Harry W. Schroeder, Jr., Anthony J. Frew, and Cornelia M. Wey. Clinical Immunology: Principles and Practice. Jr. HWS, editor: MOSBY ELSEVIER; 2008. 17 p.

78. Abolhassani H, Gharib B, Shahinpour S, Masoom SN, Havaei A, Mirminachi B, et al. Autoimmunity in patients with selective IgA deficiency. Journal of investigational

allergology & clinical immunology : official organ of the International Association of Asthmology. 2015;25(2):112-9. PubMed PMID: 25997304.

79. Jorgensen GH, Thorsteinsdottir I, Gudmundsson S, Hammarstrom L, Ludviksson BR. Familial aggregation of IgAD and autoimmunity. Clinical immunology. 2009 May;131(2):233-9. PubMed PMID: 19167929.

80. Cunningham-Rundles C, Brandeis WE, Pudifin DJ, Day NK, Good RA. Autoimmunity in selective IgA deficiency: relationship to anti-bovine protein antibodies, circulating immune complexes and clinical disease. Clinical and experimental immunology. 1981 Aug;45(2):299-304. PubMed PMID: 7318256. Pubmed Central PMCID: 1537366.

81. Wickman M, Kull I, Pershagen G, Nordvall SL. The BAMSE project: presentation of a prospective longitudinal birth cohort study. Pediatric allergy and immunology : official publication of the European Society of Pediatric Allergy and Immunology. 2002;13 Suppl 15:11-3. PubMed PMID: 12688617.

82. Magnusson PK, Almqvist C, Rahman I, Ganna A, Viktorin A, Walum H, et al. The Swedish Twin Registry: establishment of a biobank and other recent developments. Twin research and human genetics : the official journal of the International Society for Twin Studies. 2013 Feb;16(1):317-29. PubMed PMID: 23137839.

83. Frankowiack M, Kovanen RM, Repasky GA, Lim CK, Song C, Pedersen NL, et al. The higher frequency of IgA deficiency among Swedish twins is not explained by HLA haplotypes. Genes and immunity. 2015 Apr-May;16(3):199-205. PubMed PMID: 25569265.

84. Viktorin A, Frankowiack M, Padyukov L, Chang Z, Melen E, Saaf A, et al. IgA measurements in over 12 000 Swedish twins reveal sex differential heritability and regulatory locus near CD30L. Human molecular genetics. 2014 Aug 1;23(15):4177-84. PubMed PMID: 24676358. Pubmed Central PMCID: 4082371.

85. Dilthey A, Leslie S, Moutsianas L, Shen J, Cox C, Nelson MR, et al. Multi-population classical HLA type imputation. PLoS computational biology. 2013;9(2):e1002877. PubMed PMID: 23459081. Pubmed Central PMCID: 3572961.

86. Jia X, Han B, Onengut-Gumuscu S, Chen WM, Concannon PJ, Rich SS, et al. Imputing amino acid polymorphisms in human leukocyte antigens. PloS one. 2013;8(6):e64683. PubMed PMID: 23762245. Pubmed Central PMCID: 3675122.

87. Olerup O, Aldener A, Fogdell A. HLA-DQB1 and -DQA1 typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours. Tissue antigens. 1993 Mar;41(3):119-34. PubMed PMID: 8316943.

88. Cao H, Wu J, Wang Y, Jiang H, Zhang T, Liu X, et al. An integrated tool to study MHC region: accurate SNV detection and HLA genes typing in human MHC region using targeted high-throughput sequencing. PloS one. 2013;8(7):e69388. PubMed PMID: 23894464. Pubmed Central PMCID: 3722289.

89. Fang M, Abolhassani H, Lim CK, Zhang J, Hammarstrom L. Next Generation Sequencing Data Analysis in Primary Immunodeficiency Disorders - Future Directions. Journal of clinical immunology. 2016 May;36 Suppl 1:68-75. PubMed PMID: 26993986.

90. N.W. Rayner MIM. Development and Use Of a Pipeline to Generate Strand and Position Information for Common Genotyping Chips 2011. Available from: http://www.well.ox.ac.uk/~wrayner/strand/.

91. Stewart CA, Horton R, Allcock RJ, Ashurst JL, Atrazhev AM, Coggill P, et al. Complete MHC haplotype sequencing for common disease gene mapping. Genome research. 2004 Jun;14(6):1176-87. PubMed PMID: 15140828. Pubmed Central PMCID: 419796.

92. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, et al. Analysis Tool Web Services from the EMBL-EBI. Nucleic acids research. 2013 Jul;41(Web Server issue):W597-600. PubMed PMID: 23671338. Pubmed Central PMCID: 3692137.

93. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007 Nov 1;23(21):2947-8. PubMed PMID: 17846036.

94. Bakshi A, Zhu Z, Vinkhuyzen AA, Hill WD, McRae AF, Visscher PM, et al. Fast set-based association analysis using summary data from GWAS identifies novel gene loci for human complex traits. Scientific reports. 2016;6:32894. PubMed PMID: 27604177. Pubmed Central PMCID: 5015118.

95. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics. 2015 Nov 1;31(21):3555-7. PubMed PMID: 26139635. Pubmed Central PMCID: 4626747.

96. Deelen P, Bonder MJ, van der Velde KJ, Westra HJ, Winder E, Hendriksen D, et al. Genotype harmonizer: automatic strand alignment and format conversion for genotype data integration. BMC research notes. 2014;7:901. PubMed PMID: 25495213. Pubmed Central PMCID: 4307387.

97. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nature genetics. 2012 Aug;44(8):955-9. PubMed PMID: 22820512. Pubmed Central PMCID: 3696580.

98. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. American journal of human genetics. 2007 Sep;81(3):559-75. PubMed PMID: 17701901. Pubmed Central PMCID: 1950838.

99. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. Genetic epidemiology. 2000;19 Suppl 1:S36-42. PubMed PMID: 11055368.

100. Horvath S, Xu X, Laird NM. The family based association test method: strategies for studying general genotype--phenotype associations. European journal of human genetics : EJHG. 2001 Apr;9(4):301-6. PubMed PMID: 11313775.

101. Lake SL, Blacker D, Laird NM. Family-based tests of association in the presence of linkage. American journal of human genetics. 2000 Dec;67(6):1515-25. PubMed PMID: 11058432. Pubmed Central PMCID: 1287928.

102. Risch N. Assessing the role of HLA-linked and unlinked determinants of disease. American journal of human genetics. 1987 Jan;40(1):1-14. PubMed PMID: 3468804. Pubmed Central PMCID: 1684002.

103. S.A.G.E. [2012] Statistical Analysis for Genetic Epidemiology, Release 6.3 2012. Available from: http://darwin.cwru.edu

104. Gonzalez-Galarza FF, Takeshita LY, Santos EJ, Kempson F, Maia MH, da Silva AL, et al. Allele frequency net 2015 update: new features for HLA epitopes, KIR and

disease and HLA adverse drug reaction associations. Nucleic acids research. 2015 Jan;43(Database issue):D784-8. PubMed PMID: 25414323. Pubmed Central PMCID: 4383964.

105. Vorechovsky I, Zetterquist H, Paganelli R, Koskinen S, Webster AD, Bjorkander J, et al. Family and linkage study of selective IgA deficiency and common variable immunodeficiency. Clinical immunology and immunopathology. 1995 Nov;77(2):185-92. PubMed PMID: 7586726.

106. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. American journal of human genetics. 2011 Jan 7;88(1):76-82. PubMed PMID: 21167468. Pubmed Central PMCID: 3014363.

107. MacHulla HK, Schonermarck U, Schaaf A, Muller LP, Kloss C, Kruger J, et al. HLA-A, B, Cw and DRB1, DRB3/4/5, DQB1, DPB1 frequencies in German immunoglobulin A-deficient individuals. Scandinavian journal of immunology. 2000 Aug;52(2):207-11. PubMed PMID: 10931389.

108. Han S, Park J, Lee DH. Protein DHX38 is a novel inhibitor of protein phosphatase 4. Anim Cells Syst. 2015 Jul 4;19(4):236-44. PubMed PMID: WOS:000360224200002. English.

109. Wolfe D, Dudek S, Ritchie MD, Pendergrass SA. Visualizing genomic information across chromosomes with PhenoGram. BioData mining. 2013 Oct 16;6(1):18. PubMed PMID: 24131735. Pubmed Central PMCID: 4015356.

110. Furey TS, Haussler D. Integration of the cytogenetic map with the draft human genome sequence. Human molecular genetics. 2003 May 1;12(9):1037-44. PubMed PMID: 12700172.

111. Buckley RH. Clinical and immunologic features of selective IgA deficiency. Birth defects original article series. 1975;11(1):134-42. PubMed PMID: 807270.

112. Ostergaard PA. Clinical and immunological features of transient IgA deficiency in children. Clinical and experimental immunology. 1980 Jun;40(3):561-5. PubMed PMID: 7418266. Pubmed Central PMCID: 1538944.

113. Plebani A, Ugazio AG, Monafo V, Burgio GR. Clinical heterogeneity and reversibility of selective immunoglobulin A deficiency in 80 children. Lancet. 1986 Apr 12;1(8485):829-31. PubMed PMID: 2870316.

114. de Laat PC, Weemaes CM, Bakkeren JA. Immunoglobulin levels during follow-up of children with selective IgA deficiency. Scandinavian journal of immunology. 1992 Jun;35(6):719-25. PubMed PMID: 1604243.

115. Shkalim V, Monselize Y, Segal N, Zan-Bar I, Hoffer V, Garty BZ. Selective IgA deficiency in children in Israel. Journal of clinical immunology. 2010 Sep;30(5):761-5. PubMed PMID: 20571893.

116. Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. Nature reviews Genetics. 2002 Nov;3(11):872-82. PubMed PMID: 12415317.

117. Huntley CC, Stephenson RL. IgA deficiency: family studies. North Carolina medical journal. 1968 Aug;29(8):325-31. PubMed PMID: 4175121.

118. Lewkonia RM, Gairdner D, Doe WF. IgA deficiency in one of identical twins. British medical journal. 1976 Feb 7;1(6005):311-3. PubMed PMID: 942876. Pubmed Central PMCID: 1638695.

119. Ulfarsson J, Gudmundsson S, Birgisdottir B, Kjeld JM, Jensson O. Selective serum IgA deficiency in Icelanders. Frequency, family studies and Ig levels. Acta medica Scandinavica. 1982;211(6):481-7. PubMed PMID: 7113764.

120. Kouvalainen K, Moilanen I. Intrapair similarity of immunoglobulin levels in twins. Acta geneticae medicae et gemellologiae. 1987;36(4):509-15. PubMed PMID: 3454513.

121. Di Franco P, Brai M, Misiano G, Piazza AM, Giorgi G, Cossarizza A, et al. Genetic and environmental influences on serum levels of immunoglobulins and complement components in monozygotic and dizygotic twins. Journal of clinical & laboratory immunology. 1988 Sep;27(1):5-10. PubMed PMID: 3251046.

122. Alper CA, Husain Z, Larsen CE, Dubey DP, Stein R, Day C, et al. Incomplete penetrance of susceptibility genes for MHC-determined immunoglobulin deficiencies in monozygotic twins discordant for type 1 diabetes. Journal of autoimmunity. 2006 Sep;27(2):89-95. PubMed PMID: 17029885. Pubmed Central PMCID: 1810396.

123. Sun C, Molineros JE, Looger LL, Zhou XJ, Kim K, Okada Y, et al. High-density genotyping of immune-related loci identifies new SLE risk variants in individuals with Asian ancestry. Nature genetics. 2016 Mar;48(3):323-30. PubMed PMID: 26808113. Pubmed Central PMCID: 4767573.

124. Hu X, Deutsch AJ, Lenz TL, Onengut-Gumuscu S, Han B, Chen WM, et al. Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes risk. Nature genetics. 2015 Aug;47(8):898-905. PubMed PMID: 26168013. Pubmed Central PMCID: 4930791.

125. Patsopoulos NA, Barcellos LF, Hintzen RQ, Schaefer C, van Duijn CM, Noble JA, et al. Fine-mapping the genetic association of the major histocompatibility complex in multiple sclerosis: HLA and non-HLA effects. PLoS genetics. 2013 Nov;9(11):e1003926. PubMed PMID: 24278027. Pubmed Central PMCID: 3836799.

126. Zhang CE, Li Y, Wang ZX, Gao JP, Zhang XG, Zuo XB, et al. Variation at HLA-DPB1 is associated with dermatomyositis in Chinese population. The Journal of dermatology. 2016 Nov;43(11):1307-13. PubMed PMID: 27153935.

127. Walsh EC, Mather KA, Schaffner SF, Farwell L, Daly MJ, Patterson N, et al. An integrated haplotype map of the human major histocompatibility complex. American journal of human genetics. 2003 Sep;73(3):580-90. PubMed PMID: 12920676. Pubmed Central PMCID: 1180682.

128. Martin M, Mann D, Carrington M. Recombination rates across the HLA complex: use of microsatellites as a rapid screen for recombinant chromosomes. Human molecular genetics. 1995 Mar;4(3):423-8. PubMed PMID: 7795597.

129. Hemantkumar Patela YC, Titus Barnesa, Pedro Canob, Marcelo Fernandez-Vinac, Kai Caoa, editor Frequencies of recombination in the HLA class I & II regions in hematopoietic stem cell transplant patients and related donors2016: Human Immunology 77 Supplement P115, Pages 122.

130. Sullivan KA, Wolfe MA, Lopez M, Jaspan JB, Bryer-Ash M. First report of recombination between the HLA-DR and HLA-DQ loci within a family. Human immunology. 1997 Sep 15;57(1):37-43. PubMed PMID: 9438193.

131. Morris DL, Taylor KE, Fernando MM, Nititham J, Alarcon-Riquelme ME, Barcellos LF, et al. Unraveling multiple MHC gene associations with systemic lupus

erythematosus: model choice indicates a role for HLA alleles and non-HLA genes in Europeans. American journal of human genetics. 2012 Nov 2;91(5):778-93. PubMed PMID: 23084292. Pubmed Central PMCID: 3487133.

132. Fernando MM, Stevens CR, Sabeti PC, Walsh EC, McWhinnie AJ, Shah A, et al. Identification of two independent risk factors for lupus within the MHC in United Kingdom families. PLoS genetics. 2007 Nov;3(11):e192. PubMed PMID: 17997607. Pubmed Central PMCID: 2065882.

133. Goudey B, Abraham G, Kikianty E, Wang Q, Rawlinson D, Shi F, et al. Interactions within the MHC contribute to the genetic architecture of celiac disease. PloS one. 2017;12(3):e0172826. PubMed PMID: 28282431. Pubmed Central PMCID: 5345796.

134. Nah SS, Choi IY, Yoo B, Kim YG, Moon HB, Lee CK. Advanced glycation end products increases matrix metalloproteinase-1, -3, and -13, and TNF-alpha in human osteoarthritic chondrocytes. FEBS letters. 2007 May 1;581(9):1928-32. PubMed PMID: 17434489.

135. Yammani RR, Carlson CS, Bresnick AR, Loeser RF. Increase in production of matrix metalloproteinase 13 by human articular chondrocytes due to stimulation with S100A4: Role of the receptor for advanced glycation end products. Arthritis and rheumatism. 2006 Sep;54(9):2901-11. PubMed PMID: 16948116.

136. Tezuka H, Abe Y, Iwata M, Takeuchi H, Ishikawa H, Matsushita M, et al. Regulation of IgA production by naturally occurring TNF/iNOS-producing dendritic cells. Nature. 2007 Aug 23;448(7156):929-33. PubMed PMID: 17713535.

137. Sanchez-Segura A, Brieva JA, Rodriguez C. Regulation of immunoglobulin secretion by plasma cells infiltrating nasal polyps. The Laryngoscope. 2000 Jul;110(7):1183-8. PubMed PMID: 10892693.

138. Fritsche LG, Chen W, Schu M, Yaspan BL, Yu Y, Thorleifsson G, et al. Seven new loci associated with age-related macular degeneration. Nature genetics. 2013 Apr;45(4):433-9, 9e1-2. PubMed PMID: 23455636. Pubmed Central PMCID: 3739472.

139. Chen W, Stambolian D, Edwards AO, Branham KE, Othman M, Jakobsdottir J, et al. Genetic variants near TIMP3 and high-density lipoprotein-associated loci influence susceptibility to age-related macular degeneration. Proceedings of the National Academy of Sciences of the United States of America. 2010 Apr 20;107(16):7401-6. PubMed PMID: 20385819. Pubmed Central PMCID: 2867722.

140. Yu H, Yuan L, Yang Y, Ma S, Peng L, Wang Y, et al. Increased serum IgA concentration and plasmablast frequency in patients with age-related macular degeneration. Immunobiology. 2016 May;221(5):650-6. PubMed PMID: 26827241.

141. Tomer Y, Dolan LM, Kahaly G, Divers J, D'Agostino RB, Jr., Imperatore G, et al. Genome wide identification of new genes and pathways in patients with both autoimmune thyroiditis and type 1 diabetes. Journal of autoimmunity. 2015 Jun;60:32-9. PubMed PMID: 25936594. Pubmed Central PMCID: 4457545.

142. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nature genetics. 2009 Jun;41(6):703-7. PubMed PMID: 19430480. Pubmed Central PMCID: 2889014.

143. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S, et al. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. The New England journal of medicine. 2008 Feb 28;358(9):900-9. PubMed PMID: 18204098.

144. Smyth DJ, Cooper JD, Howson JM, Walker NM, Plagnol V, Stevens H, et al. PTPN22 Trp620 explains the association of chromosome 1p13 with type 1 diabetes and shows a statistical interaction with HLA class II genotypes. Diabetes. 2008 Jun;57(6):1730-7. PubMed PMID: 18305142. Epub 2008/02/27. eng.

145. Steck AK, Liu SY, McFann K, Barriga KJ, Babu SR, Eisenbarth GS, et al. Association of the PTPN22/LYP gene with type 1 diabetes. Pediatr Diabetes. 2006 Oct;7(5):274-8. PubMed PMID: 17054449. eng.

146. Bjørnvold M, Amundsen SS, Stene LC, Joner G, Dahl-Jørgensen K, Njølstad PR, et al. FOXP3 polymorphisms in type 1 diabetes and coeliac disease. J Autoimmun. 2006 Sep;27(2):140-4. PubMed PMID: 16996248. Epub 2006/09/20. eng.

147. Jacobson EM, Huber A, Tomer Y. The HLA gene complex in thyroid autoimmunity: from epidemiology to etiology. J Autoimmun. 2008 2008 Feb-Mar;30(1-2):58-62. PubMed PMID: 18178059. Pubmed Central PMCID: PMC2244911. Epub 2008/01/04. eng.

148. Hodge SE, Ban Y, Strug LJ, Greenberg DA, Davies TF, Concepcion ES, et al. Possible interaction between HLA-DRbeta1 and thyroglobulin variants in Graves' disease. Thyroid. 2006 Apr;16(4):351-5. PubMed PMID: 16646680. eng.

149. Takahashi M, Kimura A. HLA and CTLA4 polymorphisms may confer a synergistic risk in the susceptibility to Graves' disease. J Hum Genet. 2010 May;55(5):323-6. PubMed PMID: 20300120. Epub 2010/03/19. eng.

150. Kula D, Bednarczuk T, Jurecka-Lubieniecka B, Polanska J, Hasse-Lazar K, Jarzab M, et al. Interaction of HLA-DRB1 alleles with CTLA-4 in the predisposition to Graves' disease: the impact of DRB1*07. Thyroid. 2006 May;16(5):447-53. PubMed PMID: 16756466. eng.

151. Wei WH, Loh CY, Worthington J, Eyre S. Immunochip Analyses of Epistasis in Rheumatoid Arthritis Confirm Multiple Interactions within MHC and Suggest Novel Non-MHC Epistatic Signals. The Journal of rheumatology. 2016 May;43(5):839-45. PubMed PMID: 26879349.

152. Cortes A, Pulit SL, Leo PJ, Pointon JJ, Robinson PC, Weisman MH, et al. Major histocompatibility complex associations of ankylosing spondylitis are complex and involve further epistasis with ERAP1. Nature communications. 2015 May 21;6:7146. PubMed PMID: 25994336. Pubmed Central PMCID: 4443427.

153. Genetic Analysis of Psoriasis C, the Wellcome Trust Case Control C, Strange A, Capon F, Spencer CC, Knight J, et al. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. Nature genetics. 2010 Nov;42(11):985-90. PubMed PMID: 20953190. Pubmed Central PMCID: 3749730.

154. Varade J, Wang N, Lim CK, Zhang T, Zhang Y, Liu X, et al. Novel genetic loci associated HLA-B*08:01 positive myasthenia gravis. Journal of autoimmunity. 2017 Oct 13. PubMed PMID: 29037440.

155. Hughes T, Adler A, Kelly JA, Kaufman KM, Williams AH, Langefeld CD, et al. Evidence for gene-gene epistatic interactions among susceptibility loci for systemic lupus

erythematosus. Arthritis and rheumatism. 2012 Feb;64(2):485-92. PubMed PMID: 21952918. Pubmed Central PMCID: 3268866.

156. Walunas TL, Bakker CY, Bluestone JA. CTLA-4 ligation blocks CD28-dependent T cell activation. The Journal of experimental medicine. 1996 Jun 1;183(6):2541-50. PubMed PMID: 8676075. Pubmed Central PMCID: 2192609.

157. Cerutti A. The regulation of IgA class switching. Nature reviews Immunology. 2008 Jun;8(6):421-34. PubMed PMID: 18483500. Pubmed Central PMCID: 3062538.

158. Xie M, Zhang D, Zhang Y, Yang X, Su Y, Wang Y, et al. Association of genetic polymorphisms in IL-1R1 and IL-1R2 genes with IgA nephropathy in the Han Chinese population. Oncotarget. 2017 Aug 01;8(31):50673-9. PubMed PMID: 28881593. Pubmed Central PMCID: 5584187.

159. Gibson G. Hints of hidden heritability in GWAS. Nat Genet. 2010 Jul;42(7):558-60. PubMed PMID: 20581876.

160. Manolio TA. Cohort studies and the genetics of complex disease. Nat Genet. 2009 Jan;41(1):5-6. PubMed PMID: 19112455.

161. Evans DM, Spencer CC, Pointon JJ, Su Z, Harvey D, Kochan G, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. Nat Genet. 2011 Jul;43(8):761-7. PubMed PMID: 21743469. Pubmed Central PMCID: PMC3640413. Epub 2011/07/10. eng.

162. Kirino Y, Bertsias G, Ishigatsubo Y, Mizuki N, Tugal-Tutkun I, Seyahi E, et al. Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLA-B*51 and ERAP1. Nat Genet. 2013 Feb;45(2):202-7. PubMed PMID: 23291587. Pubmed Central PMCID: PMC3810947. Epub 2013/01/06. eng.

163. Strange A, Capon F, Spencer CC, Knight J, Weale ME, Allen MH, et al. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. Nat Genet. 2010 Nov;42(11):985-90. PubMed PMID: 20953190. Pubmed Central PMCID: PMC3749730. Epub 2010/10/17. eng.

164. Chitnis N, Clark PM, Kamoun M, Stolle C, Brad Johnson F, Monos DS. An Expanded Role for HLA Genes: HLA-B Encodes a microRNA that Regulates IgA and Other Immune Response Transcripts. Frontiers in immunology. 2017;8:583. PubMed PMID: 28579988. Pubmed Central PMCID: 5437213.