UNIVERSITY OF TARTU
Faculty of Science and Technology
Institute of Technology

Maksym Zarodniuk

Profiling antibody responses and their contribution to disease in APS1 rat model using high-throughput sequencing approaches

Bachelor's Thesis (12 ECTS)

Curriculum Science and Technology

Supervisor(s):
PhD student, MSc Ahto Salumets
Professor, PhD Pärt Peterson

Tartu 2022

Profiling antibody responses and their contribution to disease in APS1 rat model using high-throughput sequencing approaches

Abstract:

Autoimmune polyendocrinopathy syndrome 1 (APS1) patients with defects in central T cell tolerance collectively display a broad repertoire of high-titer autoantibodies of which antibodies to Type I interferons (IFNs) are most pathognomonic. Here we applied high-throughput immunoglobulin repertoire sequencing and single-cell transcriptomics to characterize isotype-specific antibody responses and identify regulatory and possible pathogenic mechanisms of autoantibodies in APS1 rat model. We show that increased clonality in AIRE deficiency that is dominated by the IgG2b subtype, together with skewed use of the IGHV genes, may represent autoantigen-specific signatures associated with a T_H1-dominated immune response caused by the breakdown of T cell tolerance. We also demonstrate that neutralizing antibodies to IFN-α caused pervasive downregulation of interferon-stimulated genes that are sensitive to tonic IFN at baseline. Most interestingly, this downregulation was associated with impaired activation of natural killer cells which are known to be important mediators of several immunopathologies. We propose that that this may have a broader immunosuppressive effect in APS1 as well as other immune-mediated diseases characterized by the presence of anti-IFN autoantibodies.

Keywords:

Autoimmune regulator, immune tolerance, autoantibodies, interferon activity

CERCS:

B500 Immunology, serology, transplantation

B110 Bioinformatics, medical informatics, biomathematics, biometrics

Antikeha vastuste profileerimine ning nende haigusseoselise rolli välja selgitamine APS1 roti mudelil kasutades kõrge läbilaskevõimega sekveneerimismeetodeid

Lühikokkuvõte:

1. tüüpi autoimmuunse polüendokrinopaatia sündroomiga (APS1) patsiente iseloomustab laiaulatuslik ning kõrgetasemeline autoantikehade tiiter millest kõige iseloomulikumad on tüüp I interferooni (IFN) vastased autoantikehad. Käesolevas töös rakendati kõrge läbilaskevõimega immunoglobuliinide repertuaari sekveneerimist ning üherakulist transkriptoomikat APS1 roti mudelil eesmärgiga iseloomustada autoantikehade isotüübi spetsiifilisi vastusi ning ühtlasi identifitseerida regulatoorseid ning potentsiaalselt haigusseoselisi mehhanisme. Antud töös näidati, et AIRE geeni funktsiooni kadumisega kaasneb suurenenud klonaalsus, mis on eelkõige tingitud just IgG2 alatüübi laiemast levikust ning lisaks tuvastati ka muutunud IGHV geenide kasutus. Need leiud võivad endast kujutada T-rakulise tolerantsuse kadumisest tingitud T_H1 seoselise autoantigeeni spetsiifilist mustrit. Antud töös on ka näha, et IFN-α spetsiifilised neutraliseerivad antikehad põhjustavad laiaulatusliku interferoon-stimuleeritud geenide allasurumise, iseäranis just nende geenide, mis on tundlikud ka muutustele IFN-i baastaseme juures. Huvipakkuvalt on see geeniekspressiooni allasurumine seotud NK rakkude häirunud aktivatsiooniga, mis on ise omakorda seotud mitmete immuunhaigusega. Antud töös pakutakse, et antud supressioon võib omada laiemat efekti nii APS1 puhul kui ka teistes immuunhaigustes, mida iseloomustavad IFN-i vastased autoantikehad.

Võtmesõnad:

Autoimmuunsuse regulaator, immuuntolerants, autoantikehad, interferooni aktiivsus

CERCS:

B500 Immunoloogia, seroloogia, transplantoloogia

B110 Bioinformaatika, meditsiiniinformatika, biomatemaatika, biomeetrika

TABLE OF CONTENTS

TERMS, ABBREVIATIONS AND NOTATIONS

AIRE	autoimmune regulator
APS1	autoimmune polyendocrinopathy syndrome 1
BCR	B cell receptor
$CD4^+$	cluster of differentiation 4 positive
$CD8^+$	cluster of differentiation 8 positive
CDR	complementarity determining region
CSR	class-switch recombination
GC	germinal center
IFN	interferon
Ig	immunoglobulin
IL	interleukin
ISG	interferon-stimulated gene
LN	lymph node
MHC	major histocompatibility complex
NK	natural killer cell
scRNA-seq	single-cell RNA sequencing
SHM	somatic hypermutation
SPL	spleen
TCR	T cell receptor
T_{FH}	T follicular helper cell
T_{FR}	T follicular regulatory cell
T_H1	T helper 1 cell
T_H2	T helper 2 cell
T_{reg}	T regulatory cell
UMI	unique molecular identifier

INTRODUCTION

The immune system is one of the most complex systems in the human body. Within it, billions of highly specialized immune cells orchestrate immune response against pathogens and injured host cells while at the same time preventing autoimmunity. In essence, the primary goal of the immune system is to recognize 'self' from 'non-self' and initiate appropriate responses aimed at restoring and maintaining homeostasis. Failure to do so often leads to diseases such as infections, cancer, and autoimmunity.

The autoimmune regulator gene (AIRE) plays an essential role in maintenance of T cell homeostasis by ensuring that developing thymocytes are exposed to a broad repertoire of self-antigens, allowing deletion or diversion of pathogenic immunity early on before T cells acquire their effector functions (Anderson et al., 2002; Metzger & Anderson, 2011). Reflecting its importance, AIRE deficiency results in autoimmune polyendocrinopathy syndrome 1 (APS1), characterized by organ-specific autoimmune destruction of multiple, mostly endocrine, tissues (Nagamine et al., 1997). AIRE deficiency also affects B cell tolerance, and APS1 patients collectively display a broad repertoire of high-titer autoantibodies which may be responsible for some of its clinical features (Fishman et al., 2017; Meyer et al., 2016). It has been proposed that APS1 autoantibodies may arise as a result of dysregulated germinal center reactions in which initially tolerized B cells undergo affinity maturation against self-antigens in the presence of cognate autoreactive T cells (Meyer et al., 2016). In addition, patient sera demonstrate skewed isotype frequencies, suggesting a defect in antibody class switching and an abnormal B cell memory compartment (Ronkainen et al., 2005).

In order to further characterize germinal center antibody responses in the APS1 rat model, we performed immunoglobulin repertoire sequencing of total B cells from two lymphoid organs: lymph nodes and spleen. In addition, we performed single-cell RNA sequencing of splenic immune cells to understand regulatory and possible pathogenic mechanisms of autoantibodies in the APS1 rat model. This work attempts to identify putative autoantigen-specific sequence signatures in B cell repertoires of AIRE-deficient rats, and determine the role of autoantibodies in the clinical presentation of the disease.

1 LITERATURE REVIEW

1.1 Adaptive immunity

The immune system is a very complex network of cellular and noncellular components that participate in host immunity to pathogens and damaged cells. The most primitive components of the immune system are found throughout the entire animal kingdom and comprise a system of nevertheless complex mechanisms termed innate or non-specific immunity. The innate immune system is an evolutionarily older host defense system that initiates a rapid inflammatory response that is nonspecific and, does not normally confer immune memory to the host. This is because innate immune cells which include myeloid cells and innate lymphoid cells (ILCs) rely on recognition of generic molecular features of pathogens called pathogen-associated molecular patterns (PAMPs) through a limited number of germ-line-encoded pattern recognition receptors (PRRs). PRR targets (such as lipopolysaccharides, peptidoglycans, nonmethylated CpG, and double-stranded RNA) are evolutionarily conserved which enables the innate recognition repertoire to detect a broad spectrum of PAMPs (Vivier & Malissen, 2005).

Another branch of the immune system is termed adaptive immune system and is a distinctive feature of the mammalian immunity which is believed to have originated approximately 500 million years ago in early jawed vertebrates (Flajnik & Kasahara, 2010). Adaptive immune responses are centered around B cells and T cells which, unlike innate immune cells, have antigen receptors capable of specifically recognizing a vast number of different antigens originating from infection and disease. Lymphocytes acquire the expression of antigen receptors through stochastic rearrangement of antigen receptor genes, resulting in a diverse repertoire of antigen specificities. The entire collection of antigen specificities in an individual is called the lymphocyte repertoire, and in humans it is estimated to be on the order of 10^7 to 10^9 (Abbas, 2022).

1.2 Lymphocyte repertoire generation

Each antigen receptor is made up of two arms: B cell receptors (BCRs) have light and heavy chains, and T cell receptors (TCRs) have homologous α and β chains. N-terminal domain of each antigen receptor forms the variable (V) region of the polypeptide, and is encoded by a V(D)J exon that is formed following rearrangement of variable (V), diversity (D), and

joining (J) gene segments in case of heavy and β chains, and V and J gene segments in case of light and α chains (**Figure 1**). C-terminal end of each polypeptide chain contains the constant (C) region. V, D, J and C gene segments are spatially separate within the germline loci and are brought together only in developing lymphocytes in a process called V(D)J recombination (Abbas, 2022).

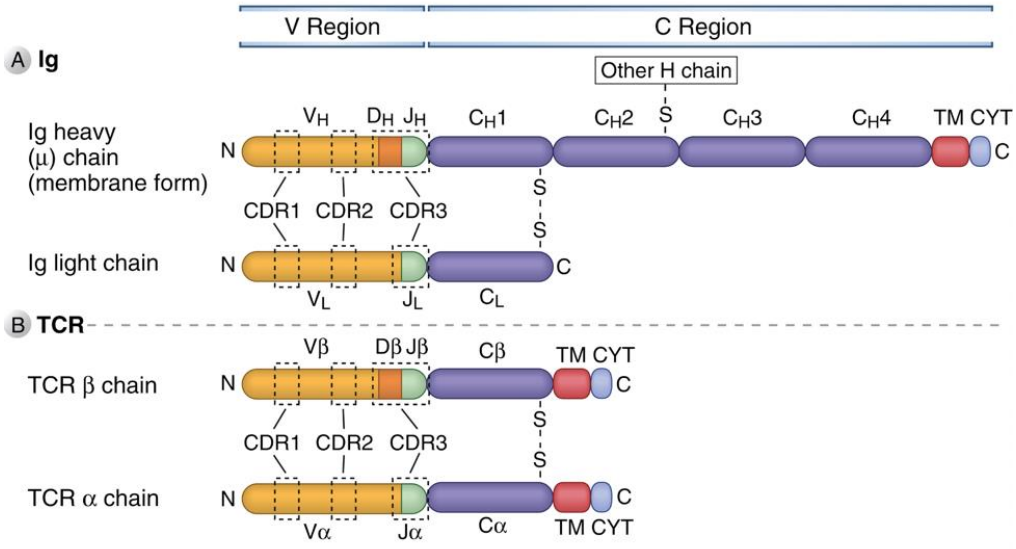

Figure 1. Structure of immunoglobulin and T cell receptor proteins. (A) Structure of immunoglobulin heavy and light chains. (**B**) Structure of TCR β and α chains. TM – transmembrane domain; CYT – cytoplasmic domain; C – carboxy terminus; N – amino terminus, CDR – complementarity determining region. Image adapted from textbook of Molecular Immunology by Abul Abbas (2022).

V(D)J recombination is achieved by chromosomal looping and cleavage of the intervening template genes that lie between the selected gene segments. Although the number of gene segments may vary across species, stochastic rearrangement of antigen receptor genes results in a combinatorial diversity of ~1,000 for each chain (Abbas, 2022). The largest contribution to the antigen receptor diversity is made by stochastic nucleotide deletions and insertions at the junctions between the newly assorted fragments, forming what is known as junctional diversity. Because of junctional diversity, antigen receptor molecules show the greatest variability at the junction between the recombined gene segments. This hypervariable region

includes the complementarity determining region 3 (CDR3), which, together with two other CDRs, constitute the antigen binding site of the receptor (**Figure 1**).

As a result of stochastic rearrangement of antigen receptor genes, immature lymphocytes often express receptors that are not functional or recognize self-antigens (or their peptides) with high affinity. To ensure generation of a functional and self-tolerant lymphocyte repertoire, cells undergo rounds of selection "centrally" in the generative lymphoid organs whereby cells expressing non-functional and strongly self-reactive receptors receive apoptosis signals. After generation and selection, lymphocytes feed the naïve repertoire where they can proliferate, mature, and differentiate as a function of signals that they receive from their local milieu. Generally speaking, naïve lymphocytes that receive stimulation by cognate antigen together with appropriate co-stimulatory signals proliferate, producing either identical progeny cells (T cells) or cells that differ by somatic point hypermutations (B cells) (**Figure 2**).

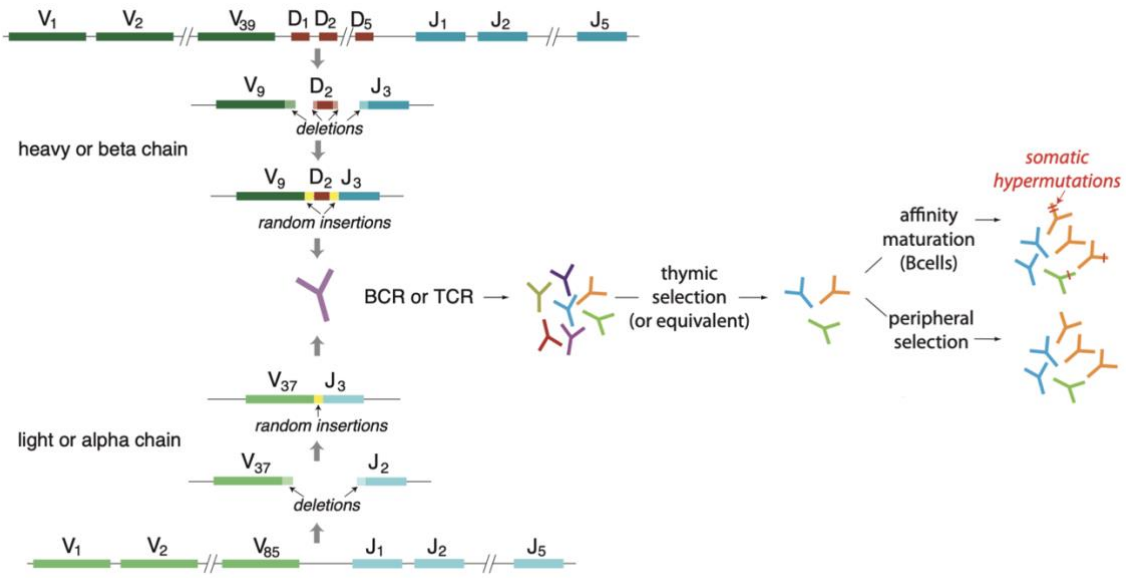

Figure 2. V(D)J recombination and generation of lymphocyte diversity. On the left, approximate organization of immunoglobulin and T cell receptor loci is shown, together with an overview of V(D)J recombination mechanism and downstream repertoire maturation. Image adapted from (Mora & Walczak, 2016).

1.3 Germinal center reaction

Germinal centers (GCs) are transient microstructures that form in the follicles of peripheral lymphoid tissues and provide a specialized microenvironment for germinal center reaction in which B cell affinity maturation and differentiation into antibody-secreting plasma cells occur. GC reaction is initiated by contemporaneous activation of $CD4^+$ helper T cells and cognate naïve B cells which brings them into close proximity at the interface of T and B cell zones where the initial antibody response develops (Chan et al., 2009; Coffey et al., 2009). Some of the activated T cells develop into T follicular helper (T_{FH}) cells and migrate into structures called follicles together with some of the activated B cells to initiate a germinal center where subsequent B cell responses occur (Nurieva et al., 2009).

B cells that have been activated by T_{FH} cells via CD40-CD40L interactions proliferate, forming the dark zone of the germinal center (Nurieva et al., 2009; Pereira et al., 2009). Within the dark zone, B cells undergo somatic hypermutation and migrate into the light zone where they encounter follicular dendritic cells displaying antigen as well as T_{FH} cells. Upon cognate interactions, B cells with the highest affinity BCRs receive CD40- and cytokine-mediated signals from the T cells which promotes their survival (Mesin et al., 2016). Following GC selection, B cells may re-enter the dark zone of the germinal center where their BCRs undergo further somatic hypermutation. This process is called affinity maturation and repeats several times until high-affinity B cells exit the germinal center as antibody-producing plasmablasts that will home to the bone marrow and will differentiate into long-lived plasma cells (Gitlin et al., 2014). Some B cells that underwent limited somatic hypermutation may leave the germinal center as memory B cells that recirculate between secondary lymphoid organs in a quiescent state until the same antigen is reencountered which triggers a potent secondary immune response (Takemori et al., 2014).

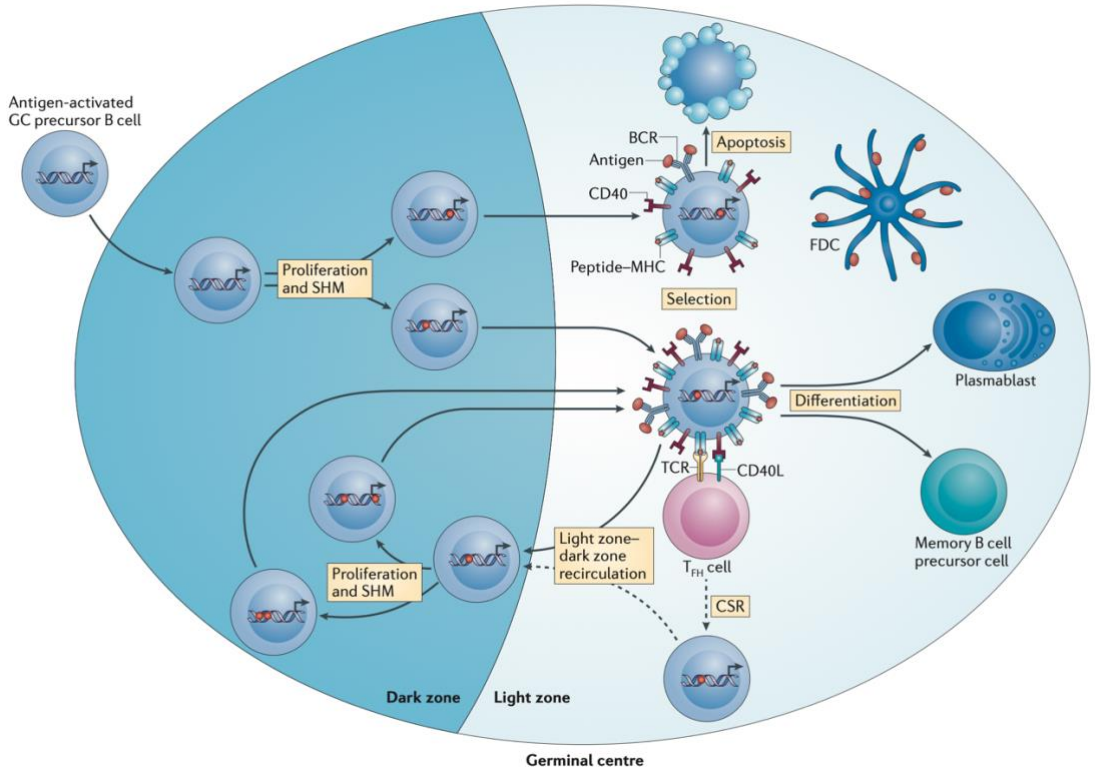

Figure 3. Germinal center reaction. Dark zone of the germinal center mostly consists of proliferating B cells that undergo somatic hypermutation – a process in which base pair changes are introduced into the variable region genes (red dots). B cells then move to the light zone in which follicular dendritic cells (FDCs) and T follicular helper (T_{FH}) cells promote the survival of B cells with the strongest affinity towards the antigen. In contrast, B cells expressing low-affinity BCRs are unable to capture sufficient antigen and undergo apoptosis. Light zone B cells may undergo immunoglobulin class-switch recombination upon receiving appropriate cytokine-mediated signals. Some B cells recirculate between the light and dark zones of the germinal center which contributes to the generation of high-affinity memory B cells and plasmablasts. CD40 – CD40 ligand; TCR – T cell receptor. Image adapted from (de Silva & Klein, 2015).

Engagement of CD40 receptor by CD40L together with cytokine stimulation by T_{FH} cells in the light zone also induces immunoglobulin class switch recombination in which the variable (VDJ) region of the Ig heavy chain is placed adjacent to one of the downstream C regions and the intervening DNA is excised (Stavnezer, 1996). After switching, the C exon closest to the VDJ exon is transcribed and translated, producing Ig proteins of the new isotype which will determine the antibody's effector function. Naïve B cells undergoing GC reaction have

the potential to switch to any isotype depending on the nature of cytokines secreted by T cells. For example, interleukin 4 (IL-4) induces preferential switching to IgG1 and IgE classes in mouse B cells, whereas IFN-γ directs switching towards IgG2a (Collins & Dunnick, 1993; Rothman et al., 1990; Stavnezer, 1996).

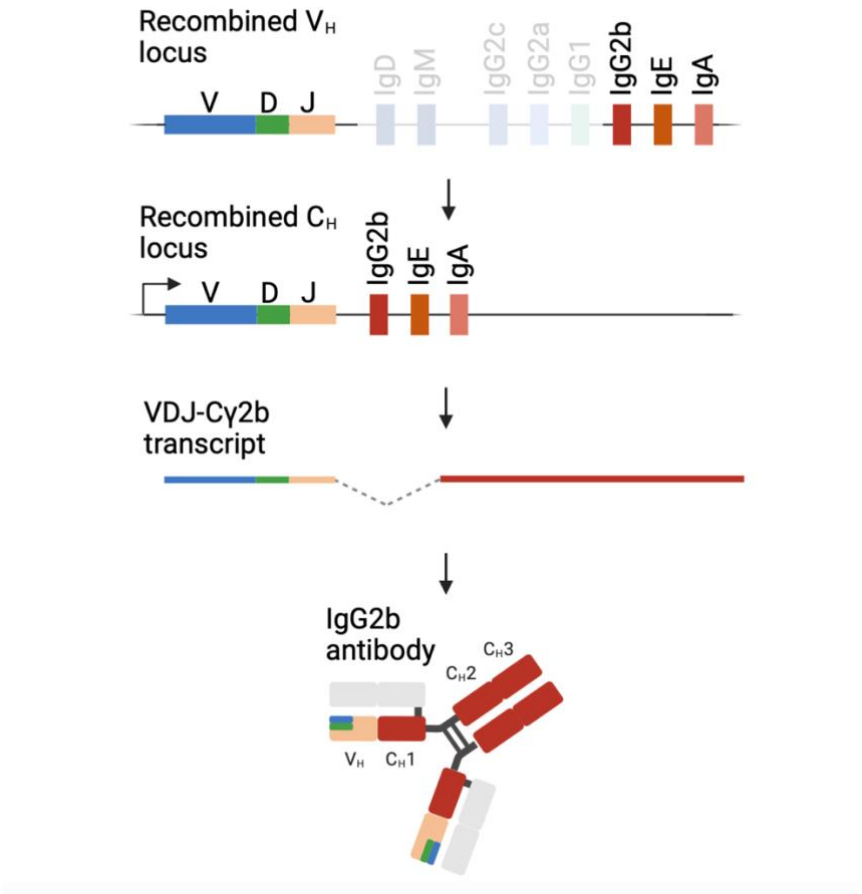

Figure 4. Immunoglobulin class switch recombination in rat. Upon receiving switch signals from other immune cells, naïve B cells in germinal centers can undergo class switch recombination in which the C_H exon corresponding to the selected class (e.g. such as IgG2b in this figure) is juxtaposed to the recombined VDJ exon and the intervening sequences are deleted (greyed out area). After rearrangement, the C_H exon most proximal to the VDJ exon is transcribed, spliced (dashed lines), and translated, generating an antibody or membrane-bound BCR (not shown) of the desired class. Image created in BioRender.

1.4 Central immune tolerance

T-cell tolerance is established centrally in the thymus which is a bilobed organ located in the central portion of the thoracic cavity. Each of the two thymic lobes is further subdivided into multiple lobules consisting of an outer cortex and an inner medulla. The removal of thymocytes whose receptors bind strongly to self-peptides presented on major histocompatibility complex (MHC) molecules occurs in the medulla which contains a specialized subset of cells called medullary thymic epithelial cells (mTECs). The mTECs express a transcriptional activator called the autoimmune regulator (AIRE) which promotes the expression of thousands of otherwise tissue-restricted genes whose products are subsequently presented to developing T cells on MHC molecules (Sansom et al., 2014). High avidity interactions between a presented antigenic peptide and its cognate TCR result in deletion of the T cell, while some of the surviving $CD4^+$ cells differentiate into tolerogenic T regulatory (T_{reg}) cells (**Figure 5**). As a result, as few as 5% of all thymocytes pass through the thymic medulla and become a part of the peripheral T-cell pool (Sinclair et al., 2013; Thomas-Vaslin et al., 2008).

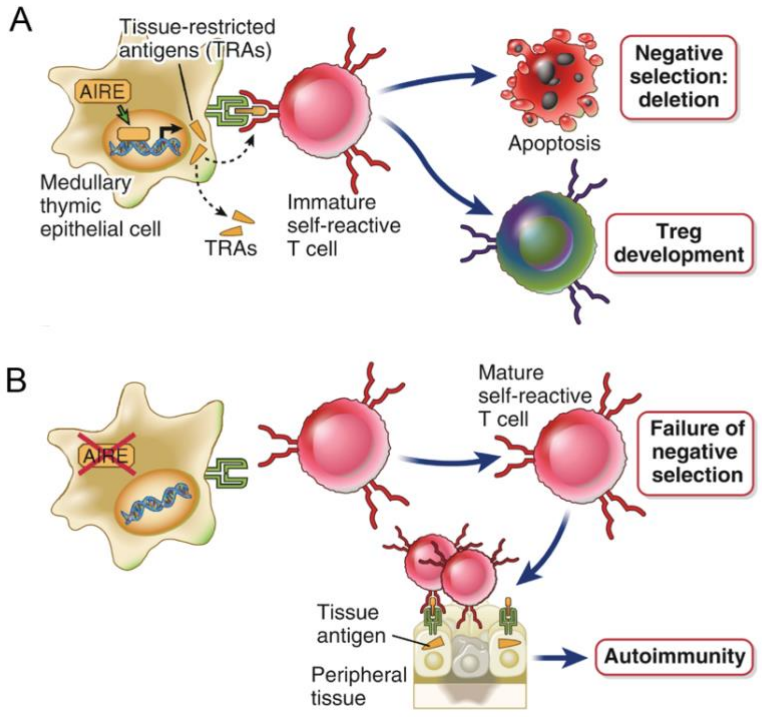

Figure 5. The function of AIRE in central T cell tolerance. **(A)** Autoimmune regulator (AIRE) is a nuclear protein that facilitates ectopic expression of tissue-restricted antigens (TRAs) in medullary thymic epithelial cells (mTECs). Immature T cells that engage strongly

with TRA peptides presented on mTECs receive apoptosis signals and die or differentiate into T regulatory (T_{reg}) cells. (**B**) In the absence of functional AIRE, self-reactive T cells escape negative selection and migrate into the periphery where they can participate in autoimmune responses. Image adapted from textbook of Molecular Immunology by Abul Abbas (2022).

1.5 Peripheral B cell tolerance

Peripheral tolerance is the second branch of immunological tolerance that takes place in the immune periphery. Although there exist a number of peripheral B cell tolerance mechanisms that do not directly rely on T cells, many B cell tolerance checkpoints are T-cell-dependent, as illustrated by a number of human diseases.

For example, accumulation of peripheral autoreactive B cells has been reported in immune dysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX) patients who have mutations in the forkhead box protein 3 *(*FOXP3*)* gene, leading to impaired differentiation of thymic T_{reg} cells (Tuure Kinnunen, 2013). The authors proposed a model in which dysfunctional T_{reg} cells act on B cells both directly by failing to suppress them, as well as indirectly by failing to control the activation of $CD4^+$ helper and T_{FH} cells, leading to selection of autoreactive B cells. Compromised B cell tolerance as a result of T_{reg} dysfunction has also been reported in rheumatoid arthritis (RA) patients whose T_{reg} display an anergic phenotype and fail to secrete anti-inflammatory cytokines transforming growth factor β (TGF-β) and IL-10, resulting in accumulation of autoreactive naïve B cells in the periphery (Nadkarni et al., 2007).

Another line of evidence suggests that T/B-cell interactions may play an important role in tolerizing autoreactive B cells in the periphery. For example, bare lymphocyte syndrome (BLS) patients who are deficient for MHC class II and CD40L-deficient patients have a high proportion of autoreactive mature naïve B cells, suggesting that pMHC/TCR and CD40/CD40L interactions may be involved in regulating B-cell peripheral tolerance (Meffre & Wardemann, 2008). Additionally, patients with thymic malignancies called thymomas often develop autoimmune disorders such as myasthenia gravis (MG). It has been previously shown that neoplasms of cortical TECs in MG patients often under-express MHC class II, resulting in impaired thymopoiesis (Kadota et al., 2000). In one study, broadly-reactive neutralizing autoantibodies against type I interferons were found in ~70% of thymoma-

associated MG patients (Meager et al., 2003), raising a possibility that the breakdown of B cell tolerance may be caused by the presence of self-reactive T cells in the peripheral repertoire as a result of thymic hyperplasia.

Autoreactive B cell clones can also be generated as an unintended consequence of somatic mutation during GC reaction, provided that cognate autoreactive T_{FH} cells are also present. This is exemplified by anti-desmoglein-3 antibodies in autoimmune pemphigus. One study found that binding to desmoglein-3 was lost when somatic mutations were reverted to the germline sequence (di Zenzo et al., 2012). In yet another disease called pulmonary alveolar proteinosis, characterized by the presence of anti-granulocyte macrophage-colony stimulating factor (GM-CSF) antibodies, binding to GM-CSF was found to be dependent on somatic mutations (Piccoli et al., 2015). This suggests that in both diseases autoantibodies arise as a result of dysregulated GC reactions in which initially tolerized B cells undergo T-cell-dependent affinity maturation against self-antigens.

1.6 Autoimmune polyendocrynopathy syndrome 1

Autoimmune polyendocrynopathy syndrome type-1 (APS1) is a rare monogenic disease caused by mutations in the Autoimmune regulator (AIRE) gene that are inherited in autosomal recessive manner. AIRE deficiency results in accumulation of peripheral autoreactive T cells that target multiple, most commonly ectodermal and endocrine tissues (Peterson et al., 2004). Although the disease is very heterogeneous and displays variable penetrance, APS1 patients commonly have at least two of the diagnostic 'APS1 triad': chronic mucocutaneous candidiasis (CMC), adreno-cortical failure (also known as Addison's disease) and autoimmune hypoparathyroidism (Peterson et al., 2004). One of the most clinically important features of APS1 is the presence of a broad repertoire of high titer autoantibodies targeting multiple self-antigens (Fishman et al., 2017).

APS1 patients display high inter-individual variation in their autoantigen profiles. In one study, 81 patients collectively harbored antibodies against 3,731 distinct self-proteins, while any one patient displayed autoreactivities toward 90-100 self-proteins, including a common set of 10-15 "public" (i.e. shared) specificities (Meyer et al., 2016). Interestingly, the authors found low-level autoreactivity against about 400 self-proteins in 21 healthy control sera, indicating a low-level leakiness of central immune tolerance mechanisms even in healthy individuals. However, high-titer, high-affinity serum autoantibodies were the defining

feature of APS1 patients. In another study, APS1 autoantibodies were found to be directed against two distinct subgroups of antigens: tissue-restricted antigens, many of which are expressed ectopically in mTECs, and non-tissue-restricted autoantigens enriched in lymphoid tissue proteins (Fishman et al., 2017).

The latter includes strong anti-cytokine reactivities, including T_H17-associated cytokines such as IL-17A, IL-17F, and IL-22 which have been associated with CMC (Kisand et al., 2010). Autoantibodies with the highest reported prevalence and titers target type I interferons (particularly IFN-α and IFN-ω) and have been found in nearly all of the patients studied, regardless of their clinical phenotype (Fishman et al., 2017; Meager et al., 2008; Meyer et al., 2016). These antibodies are some of the highest affinity antibodies described with near-femtomolar dissociation constants and have a higher neutralization potency than commercially available monoclonal antibodies such as sifalimumab (Meyer et al., 2016). Because such properties are rare among antibodies induced by immunization, studying etiology of type I IFN autoantibodies can not only give insights into the nature of B cell autoimmunity in APS1 patients, but also instruct the design of next generation vaccines and antibody therapeutics.

To this end, Meyer et al. (2016) found that autoantibodies targeting type I IFNs, IL22, and IL17 were highly mutated compared to their germline (i.e. unmutated) counterparts. Their germline counterparts, however, did not engage with the target antigens, suggesting that antibodies developed during affinity maturation from non-self-reactive progenitors, similarly to anti-DSG3 antibodies in autoimmune pemphigus or GM-CSF antibodies in pulmonary alveolar proteinosis.

1.7 APS1 *in vivo* models

Since identification of AIRE as the cause of APS1, several mouse models of AIRE deficiency have been developed. Despite having provided important insights into the functional relationship between *Aire* and disease mechanisms, mouse models did not accurately reproduce key phenotypical features of APS1 patients, such as severe multi-organ autoimmunity and presence of high-titer autoantibodies. Even $Aire^{-/-}$ mice with a highly autoimmune-prone non-obese diabetic (NOD) genetic background failed to develop alopecia, vitiligo, and nail dystrophy, in contrast to APS1 patients (Gavanescu et al., 2007).

A recently described AIRE-deficient rat model, however, presents a larger array of autoimmune manifestations and more accurately recapitulates the severe clinical features of the human disease (Ossart et al., 2018). At 6 months of age, *Aire*$^{-/-}$ rats start to develop skin disorders including alopecia, vitiligo, and nail dystrophy (**Figure 4A**). Several peripheral organs, including pancreas, liver, kidney, and lung show extensive lymphocytic infiltration and severe autoimmune lesions (Ossart et al., 2018).

Importantly, AIRE-deficient rats develop APS1-specific autoantibodies targeting organs (**Figures 4B-D**) and cytokines (**Figure 4C**) at levels comparable to human APS1 disease, making it an appropriate model to study antibody responses in the context of APS1. It is known that in *Aire*$^{-/-}$ rats multiple organs including intestines, kidney, thyroid, liver, testis, adrenal gland, and exocrine pancreas are targeted by autoantibodies (Ossart et al., 2018). In addition, high-affinity neutralizing autoantibodies against type I IFNs are also present and become detectable at 4-5 months and reach maximum at month 10 (**Figure 4D**). In terms of antibody isotypes, sera of *Aire*$^{-/-}$ rats contain higher levels of IgG1, IgM, and IgA, suggesting a defect in Ig class switching in germinal centers (Ossart et al., 2018).

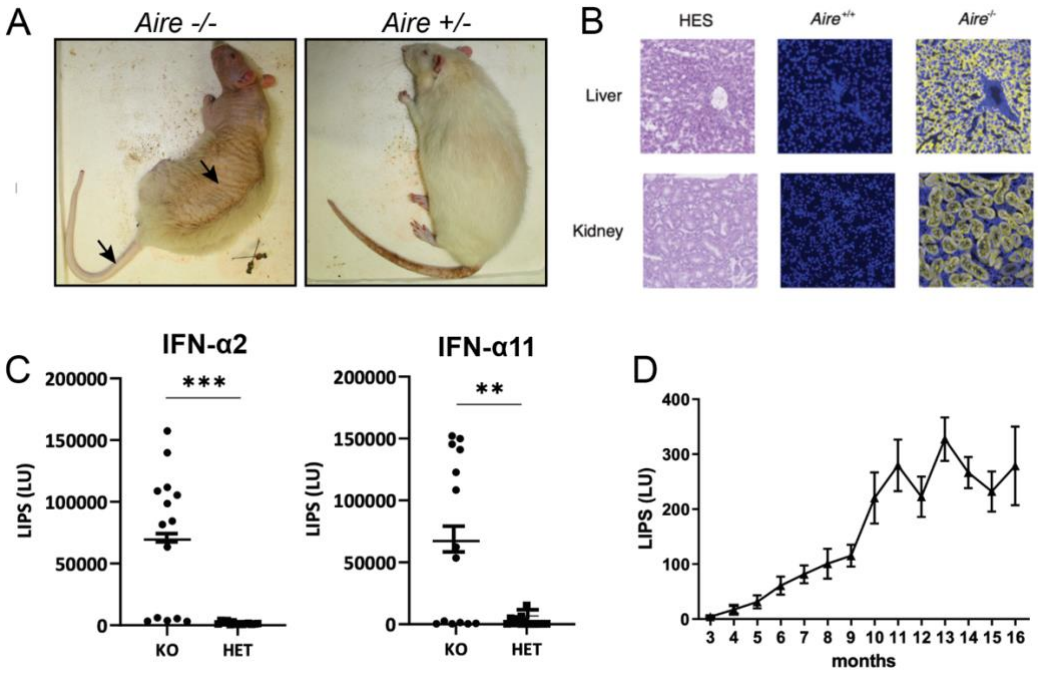

Figure 6. AIRE-deficient rats have a severe autoimmune phenotype similar to that of APS1 patients. (A) Photographs showing visual features of >1-y-old *Aire*$^{-/-}$ and *Aire*$^{+/-}$ rats. Arrows indicate alopecia and vitiligo. (**B**) Examples of tissues targeted by autoantibodies. Tissue sections of *IgM*$^{-/-}$ rats stained with HES, DAPI and sera from either *Aire*$^{+/+}$ (middle

column) or *Aire*$^{-/-}$ (right column) animals. Adapted from (Ossart et al., 2018). (**C**) Antibodies to IFN-α2 and IFN-α11 from 4-5 year-old animals were tested using luciferase immunoprecipitation system (LIPS). Each dot represents a luminescence value for one individual. Error bars indicate SEM. *** $p \leq 0.001$, ** $p \leq 0.01$, Student's t-test. **D**, Anti-IFN-α2 antibody dynamics in *Aire*$^{-/-}$ animals. Error bars indicate SEM. Figures C and D are courtesy of our lab staff.

2 THE AIMS OF THE THESIS

Most studies into B cell responses of APS1 patients have thus far been performed on the B-cell fraction of peripheral blood mononuclear cells (PBMCs), since peripheral blood is the most ethical and practical source of B cells (Raybould et al., 2021). However, peripheral blood is not a lymphoid organ where B cell responses take place, and as a result, the true disease-associated signal may be diluted (Blanco et al., 2018; Briney et al., 2014). Because obtaining samples directly from lymphoid organs by studying APS1 patients *post mortem* is often difficult, *in vivo* models of APS1 represent an ideal tool for studying antibody responses in AIRE deficiency. Therefore, in this work we used a novel physiologically relevant APS1 rat model to characterize its B cell repertoires from lymphoid tissues using immunoglobulin repertoire sequencing and single-cell transcriptomics. In this work we aimed to:

- identify putative autoantigen-specific signatures in B-cell repertoires of AIRE-deficient rats, including clonal expansions, V gene segment biases and antigen-driven convergence;
- describe the antibody class-switching landscape in APS1 rats;
- identify immunomodulatory mechanisms of autoantibodies in APS1 rats.

3 EXPERIMENTAL PART

3.1 MATERIALS AND METHODS

3.1.1 Materials

- Agarose
- Agencourt AMPure XP (Beckman Coulter, cat. no. A63881)
- Chloroform (Sigma-Aldrich, cat. no. C2432-25ML)
- DNA Clean & Concentrator-5 (Zymo Research, cat. no. D4014)
- dNTPs (Solis BioDyne)
- Ethidium bromide (Naxo)
- FIREPol DNA Polymerase (Solis BioDyne, cat. no. 01-01-00500)
- $MgCl_2$ (Naxo)
- Nextera XT DNA Library Preparation Kit (cat. no. FC-131-1096)
- Oligonucleotides
- Q5 Hot-start High-Fidelity DNA Polymerase (New England BioLabs, cat. no. M0493S)
- RNAse-free water
- RNeasy Mini Kit (Qiagen, cat. No. 74106)
- SMARTscribe reverse transcriptase (Clontech, cat. no. 639536)
- Tris-HCl (Sigma-Aldrich)
- Trizol Reagent (Invitrogen, cat. no. 15596-026)
- Uracil DNA glycosylase (New England BioLabs, cat. no. M0280S)
- 100 bp DNA Ladder Ready to Load (Solis BioDyne, cat. no. 07-11-0000S)
- 10x Yellow Buffer (Naxo)
- 96% (vol/vol) ethanol

3.1.2 Animals

We used *Aire*-mutant Sprague Dawley (SPD) rats from Guilloneau lab (Ossart et al., 2018). Homozygous $Aire^{-/-}$ and heterozygous $Aire^{+/-}$ were bred and maintained at the animal facility of the Institute of Molecular and Cell Biology at the University of Tartu. Because APS1 is inherited recessively and $Aire^{+/-}$ rats do not present with autoimmunity, we used $Aire^{+/-}$ rats as healthy controls.

3.1.3 Total RNA extraction

In order to minimize cell sampling bias and to average Ig mRNA expression levels between many cells, the initial amount of tissue specimen used was as high as possible. First, tissues were homogenized with a pestle and incubated in 1 ml of Trizol reagent for 5 min at room temperature. 0.2 ml of chloroform was added into the samples and the tubes were vortexed for 15 seconds and incubated for 5 more minutes at room temperature. Next, the samples were centrifuged at 12,000 rpm for 15 min at 4°C. After centrifugation, the mixture separated into a red phenol-chloroform phase at the bottom and a colorless aqueous phase at the top which contained RNA. The aqueous phase was transferred into a fresh tube and RNA was then precipitated and washed using RNeasy Mini Kit according to manufacturer's instructions. Agilent Bioanalyzer was used to verify the quality and concentration of each extraction product.

3.1.4 First strand cDNA synthesis with

cDNA synthesis starts with a set of C-region-specific primers to generate a specific single-stranded cDNA product together with untemplated C nucleotides and a unique molecular identifier (UMI) at the 3' end of the cDNA. In downstream data analysis, UMIs will be used to assign each sequencing read to a particular template molecule, allowing us to efficiently normalize the data and distinguish somatic mutations from PCR and sequencing errors (Turchaninova et al., 2016). The reaction (mix 1) is prepared in a sterile 0.2 ml reaction tube for a final volume of 8 µl (**Table 1**).

The reactions were placed into a thermal cycler and incubated for 2 min at 70 °C. The temperature was then decreased to 42 °C to anneal the synthesis primers for 1-3 min. The tubes were kept in the thermal cycler at 42 °C for no longer than 15 min while preparing mix 2 (**Table 2**). A heated lid was used to maintain a homogeneous state and temperature.

12.5 µl of mix 2 was then added to mix 1. The reaction mix was gently pipetted and incubated for 60 min at 42 °C. Next, 1 µl of uracil DNA glycosylase was added to remove residual template switch adaptors and incubated for 40 min at 37 °C. Following incubation, cDNA was purified using Zymo DNA Clean & Concentrator kit. The first PCR reaction was carried out the same day to avoid losing the material.

Table 1. cDNA synthesis rection with template switch. Mix 1 reaction components and their final concentrations

Component	**Amount (µl)**	**Final amount/concentration[1]**
RNA	2-6	Up to 700 ng
cDNA synthesis primer mix (10 µM each)	2	1 µM for each primer
RNAse-free water	0-4	

Table 2. cDNA synthesis rection with template switch. Mix 2 reaction components and their final concentrations

Component	**Amount (µl)**	**Final concentration**
5× First-strand buffer Clontech (Takara)	4	1×
DTT (20 mM, from SMARTscribe)	2	2 mM
5'-Template switch adaptor SmartNNNa (10 µM)[2]	2	1 µM
dNTP solution (10 mM each)	2	1 mM each
Ribolock 40 U/µl	0.5	
SMARTscribe Reverse Transcriptase (10×, Clontech)	2	10 U/µl

[1] Final concentration refers to the concentration following the addition of mix 2.

[2] Final concentration refers to the concentration following the addition of mix 2 to mix 1.

3.1.5 First PCR amplification

The first amplification is a multiplex PCR reaction which is semi-nested at the 3' end. A universal 3' adaptor is introduced during the first PCR amplification in order to avoid the need for multiplexing in the next step. The reaction was prepared in a sterile thin-walled 0.2-ml reaction tube in a final volume of 50 µl (**Table 3**). PCR was performed using parameters given in **Table 4**.

Table 3. First PCR. PCR reaction components and their final concentrations.

Component	**Amount (µl)**	**Final concentration**
Nuclease-free water	31	——
Q5 polymerase buffer (5x, NEB)	10	1×
dNTP mix (10 mM each)	1	0.2 µM each
Primer Rep_M1SS (10 µM)	2,5	0.5 µM
Primer rIGX_r3 (10 µM)	2,5	0.5 µM
Q5 Polymerase (NEB)	0,5	0.02 U/µl
First-strand cDNA	2,5	——

Table 4. First PCR program. PCR cycles used to amplify cDNA.

Cycle	**Denature**	**Anneal**	**Extend**
1	95 °C for 1 min 30s	——	——
18	95 °C for 10 s	60 °C for 20 s	72 °C for 40 s
1	——	——	72 °C for 4 min

PCR products were then purified using Zymo DNA Clean & Concentrator kit. The purified products were stored at 4 °C overnight.

3.1.6 Second PCR amplification

The second amplification is a nested (step-out) PCR. Second PCR mix was prepared in a sterile thin-walled 0.2-ml reaction tube in a final volume of 50 µl (**Table 5**). PCR was performed using parameters given in **Table 6**.

Table 5. Second PCR. PCR reaction components and their final concentrations.

Component	**Amount (µl)**	**Final concentration**
Nuclease-free water	31.5	——
Q5 polymerase buffer (5x, NEB)	10	1×
dNTP mix (10 mM each)	1	200 µM
Primer Fwd_OH_M1S (10 µM)	2.5	0.5 µM
Primer RevOH_Z2 (10 µM)	2.5	0.5 µM
Q5 Polymerase (NEB)	0,5	0.02 U/µl
Purified first PCR product	2	——

Table 6. Second PCR program. PCR cycles used to amplify first PCR product.

Cycle	**Denature**	**Anneal**	**Extend**
1	95 °C for 1 min 30s	——	——
18	95 °C for 10 s	60 °C for 20 s	72 °C for 40 s
1	——	——	72 °C for 4 min

Products from second PCR were purified using Zymo DNA Clean & Concentrator kit within an hour after amplification in order to preserve the integrity of the PCR product. Concentration of each PCR product was determined using Qubit Fluorometer according to manufacturer's instructions.

3.1.7 Sequencing library control reaction

The presence of the necessary fragment in obtained PCR products was verified by performing a sequencing library control reaction. The reaction was prepared in a sterile thin-walled 0.2-ml reaction tube in a final volume of 25 µl (**Table 7**). PCR was performed using parameters given in **Table 8**.

The products were run on 1.5% agarose gel. A sharp, visible band at 350 bp indicated high quality of the product. The products were stored at 4 °C overnight or at 20 °C for extended storage.

Table 7. Sequencing library control PCR. PCR reaction components and their final concentrations.

Component	**Amount (µl)**	**Final concentration**
Nuclease-free water	16.1	——
Yellow buffer 10x	2.5	1×
$MgCl_2$ (25 mM)	1.5	1.5 mM
dNTPs (2 mM)	1.5	0.1 µM
Primer VLFwd 1 (10 µM)	1.0	0.5 µM
Primer rIgG C1 rev (10 µM)	1.0	0.5 µM
FIREPol DNA Polymerase	0.4	0.08 U/µl
DNA template	1.0	——

Table 8. Sequencing library control PCR program. PCR cycles used to amplify the VDJ sequence fragment.

Cycle	Denature	Anneal	Extend
1	95 °C for 15 min	—	—
15	95 °C for 30 s	60 °C for 30 s	72 °C for 30 s
1	—	—	72 °C for 5 min

3.1.8 Sequencing library preparation

Concentration of each PCR product was determined using Qubit Fluorometer and the samples were diluted to equal concentration of 10 ng/µl. To generate Ig sequencing libraries with Illumina adaptors and sample barcodes on both 5' and 3' ends, indexing PCR is performed. Indexing PCR mix was prepared in a sterile thin-walled 0.2-ml reaction tube in a final volume of 15 µl (**Table 9**). PCR was performed using parameters provided in **Table 10**.

After indexing PCR, the libraries were purified with Agencourt AMPure XP purification beads. First, AmpX purification beads were allowed to equilibrate to room temperature and were vigorously shaken to resuspend any magnetic particles that may have settled to the bottom. Next, 27 µl of AMPure XP beads were added to 15 µl of each sample. The resulting suspension was pipette-mixed 10 times and the samples were allowed to incubate for 5 minutes at room temperature for maximum recovery. Then the reaction plate was placed onto a magnetic stand for 2 minutes until the supernatant has cleared. With the reaction plate still on the stand, the supernatant was carefully removed and discarded. 200 µl of freshly made 70 % ethanol was then added to each well of the reaction. After 30 seconds of incubation at room temperature, the entire ethanol wash was aspirated and discarded. A total of two washes were performed until the beads were dried. The reaction plate was then removed from the magnetic stand and 13.75 µl of 10 mM Tris-HCl (pH 8.0) was added to each well of the reaction plate and pipetted-mixed. After 2 minutes of incubation, the reaction plate was placed onto the magnetic stand for 1 minute to separate the beads from the solution. 12.5 µl of the eluate was transferred to a sterile 0.2-ml tube and 1 µl aliquots

were taken for quality control with Agilent Bioanalyzer. All samples with Illumina adapters were pooled together in equal volumes (2 µl).

Table 9. Indexing PCR. PCR reaction components and their final concentrations.

Component	**Amount (µl)**	**Final concentration**
Purified second PCR product	2,5	—
KAPA HiFi HotStart PCR Mix (2x)	7,5	1x
Nextera XT 1 Index (N7xxx)	2,5	—
Nextera XT 2 Index (E5xxx)	2,5	—

Table 10. Second PCR program. PCR cycles used to introduce sample-barcoded Illumina adaptors.

Cycle	**Denature**	**Anneal**	**Extend**
1	95 °C for 3 min	—	—
7	95 °C for 30 s	55 °C for 30 s	72 °C for 30 s
1	—	—	72 °C for 5 min

3.1.9 Sequencing and data preprocessing

High-throughput sequencing was perfomed on the Illumina MiSeq platform using symmetric 300 bp paired-end sequencing with dual-index configuration. Sequencing reads were preprocessed using the MiNNN (v10.1) software package with default settings to perform UMI-based error correction. MiXCR (v3.0.13) was used to assemble CDR3 sequences with default parameters except the default aligner was replaced with a newer kAligner2 which is specifically optimized for analysis of B-cell data with long gaps and insertions. Since clustering was performed manually in the next step, *-OcloneClusteringParameters* was set to “null”.

3.1.10 B cell lineage reconstruction

To cluster B-cell clones into lineages, we modified a previously described method to suit our data(Horns et al., 2016). Within each sample, sequences sharing the same V-J combination and CDR3 length were grouped, as these likely represent sequences originating from the same recombination event. To reduce false positives due to antigen-driven convergence, groups were further subclustered by performing average linkage clustering using scaled Levenshtein distance. The resulting trees were cut at a height of 0.85, resulting in groups of sequences sharing at least 85% sequence identity across CDR3 (i.e. sequences below this similarity threshold were grouped into separate lineages).

To choose the optimal cutoff for identifying B-cell lineages, we examined the distribution of pairwise sequence identity within groups of CDR3 sequences (**Figure S2A**). We observed a bimodal distribution corresponding to sequences with a highly similar neighbor (>85% identity), corresponding to clonally related sequences with somatic mutations, and sequences whose nearest neighbor was substantially dissimilar (<85% identity), corresponding to phylogenetically unrelated sequences. Thus, we determined the optimal cutoff that discriminated the two modes of the distribution.

The resulting groups of CDR3 sequences were used to infer lineage trees. First, germline sequence was reconstructed for each group by aligning V and J germline genes and introducing a gap corresponding to the D gene and untemplated nucleotides. Next, a matrix representation of possible ancestor-child relations was constructed by computing Hamming distance between each pair of sequences and pruning the edges that violate constant region locus architecture. The graph was rooted on the germline sequence and the minimum evolution tree for each lineage was obtained using Edmond's algorithm which finds optimum branching in a directed graph.

3.1.11 Repertoire diversity and clonality

To assess the clonality of Ig repertoires, we used a number of different diversity indices. First, we tried to assess the unevenness of the number of unique VDJ sequences per clonotype using the Simpson diversity index (Simpson, 1949). Simpson index, similarly to other diversity measures, is related to a generalized family of diversity measures called Rényi entropy which is parametrized by q and can be expressed as:

$$^{q}H = \left(-\ln \sum_{i=1}^{S} p_i^q \right) / (q-1)$$

where S is the number of species, p_i is the relative frequency of the i-th species, and q is called order. However, because entropy only gives the uncertainty in the species identity in the sample and not the number of species, it is often converted to true diversity by taking exponent of ^{q}H:

$$^{q}D = e^{\,^{q}H} = \left(\sum_{i=1}^{S} p_i^q \right)^{1/(1-q)}$$

which defines a generalized class of diversity indices called Hill diversities (1974). When $q = 2$,

$$^{2}D = 1 / \left(\sum_{i=1}^{S} p_i^2 \right)$$

This is called inverse Simpson index or true diversity of order 2. Simpson index is equal to

$$^{2}D = \sum_{i=1}^{S} p_i^2$$

We also used Chao1 which is another widely used diversity estimator. In contrast to Simpson index, Chao1 gives more weight to the low abundance species and is an estimator of species richness rather than species evenness (Chao, 1984). Chao1 estimator can be expressed as

$$S_{Chao1} = S_{obs} + \frac{F_1(F_1 - 1)}{2(F_2 + 1)}$$

Where F_1 and F_2 are the count of singletons and doubletons (i.e. clonotypes represented by one or two unique CDR3 sequences), respectively, and S_{obs} is the number of observed species (clonotypes).

D50 is a recently developed metric for immune diversity estimation. It calculates the minimum number of distinct clonotypes amounting to greater than or equal to 50% of total sequencing data.

Because at low sample coverage all diversity metrics depend on the analysis depth, we normalized samples to the same sequencing depth by bootstrapping all diversity measures 100 times to get a normalized diversity estimate.

3.1.12 Quantification of class switch events

Class switch events were identified by traversing the B-cell lineage trees and searching for ancestor-child pairs having different classes. Class-switch events that were not detected in all of the samples were filtered out. To study the class switch landscapes, we used the following metrics that were used in a previous study of human Ig repertoires (Horns et al., 2016). $P_{relative}$ defines relative switch frequency and is equal to the number of switches from A class to B class, normalized by the total number of class-switch events (i.e. total number of ancestor-child pairs of different classes):

$$P_{relative} = \frac{A \to B}{X \to X}$$

$P_{destination}$ defines the probability that a cell switching from class A will choose class B as the destination, A denoting ancestor and B denoting child class:

$$P_{destination} = \frac{A \to B}{A \to X}$$

$P_{arrival}$ defines the probability that a cell of a given downstream class B will originated from the ancestral class A:

$$P_{arrival} = \frac{A \to B}{X \to B}$$

In addition to $P_{relative}$, $P_{destination}$, and $P_{arrival}$ which describe class switch events in which ancestor and child sequences must be of different classes, we also calculated "transition probabilities" which allowed for *A* and *B* to be the of the same class.

3.1.13 Single-cell RNA sequencing data analysis

Single-cell RNA sequencing data were mapped to the *Rattus norvegicus* reference genome (Rnor_6.0) and quantified using the Cell Ranger software suite (Version 6.0.1, 10x Genomics Inc.). *Scrublet* algorithm was used to identify doublets which were removed from downstream analysis (Wolock et al., 2019), and the *default* set of filters was used to exclude cells as outliers, as described previously (Germain et al., 2020). Seurat's *sctransform* algorithm was used to normalize and scale the data with default parameters (Hafemeister & Satija, 2019), and individual samples were then harmonized using Seurat's integration pipeline (Stuart et al., 2019).

Pseudotemporal ordering of cells was inferred using *Slingshot* (Street et al., 2018), and genes whose expression changed along the trajectory were identified using *tradeSeq* which fits a negative binomial generalized additive model to expression data to obtain a smoothed expression estimate and tests the null hypothesis that gene expression is not a function of pseudotime (van den Berge et al., 2020).

Type I interferon score was calculated as described previously (Nezos et al., 2015). For each cell type, the mean and standard deviation of expression of a set of tonic-sensitive interferon-stimulated genes in control rats were used to standardize expression levels of corresponding genes in each *Aire*$^{-/-}$ rat. The standardized expression levels were subsequently summed for each cell type in *Aire*$^{-/-}$ rats to provide an "IFN score".

3.2 RESULTS

3.2.1 Increased clonality and skewed isotype usage

To study affinity maturation and class switching of B cells in the APS1 rat model, we performed bulk Ig repertoire sequencing of total B cells from lymph nodes and spleens of 13 *Aire*$^{-/-}$ and 8 *Aire*$^{+/-}$ animals. In order to resolve Ig isotypes and subtypes, we modified an existing Ig profiling protocol by Turchaninova and colleagues to include a longer stretch of the C-terminal region which allowed us to retain more information on Ig isotype (Turchaninova et al., 2016).

The resulting Ig-seq dataset consisted of 42 samples, with an average depth of roughly 420,000 reads per sample and average coverage of 1.41 reads per UMI. Filtering, UMI-based correction of PCR and sequencing errors was performed using the MiNNN software and resulting consensus sequences were mapped and assembled into clones using the MiXCR software package (**Figure S1A-B**). In order to group sequences into clonal lineages (clonotypes), we first examined the distribution of pairwise sequence identity within groups of CDR3 sequences sharing the same length and V and J gene combinations and determined the optimal cutoff that discriminated the two modes of this distribution (**Figure S2B, Methods 3.1.10**). Clonotypes were called within each group by performing average linkage clustering with a cutoff of 85% sequence identity across CDR3.

We first examined individual isotype frequencies in the data by computing the percentage of unique VDJ sequences per isotype, thus controlling for differential RNA per cell. *Aire*$^{-/-}$ animals had increased frequency of IgG2b subtype in lymph node ($P < 0.0001$) and spleen ($P < 0.05$), whereas decreased frequency of IgM clones ($P < 0.05$) was characteristic only of lymph node Ig repertoires (**Figure 1A**). Interestingly, IgG2b subtype frequency in lymph nodes seemed to be more strongly correlated with the age of *Aire*$^{-/-}$, but not *Aire*$^{+/-}$ animals ($P = 0.058$, permutation test), raising a possibility that IgG2b antibodies may be enriched in anti-IFN specificities which also become more prevalent with age (**Figures 1B and S1B**).

We next hypothesized that skewed isotype use could be due to dysregulated germinal center reaction in which autoreactive T cells drive selective expansion of B-cell clones with increased affinity for self-antigens. To test for the presence of clonal expansions, we examined diversities of Ig repertoires as described previously (Bashford-Rogers et al., 2019). Briefly, we analyzed clonotype distributions using inverse Simpson index which provides a measure of B-cell clonal expansion and BCR diversification through SHM (**Methods 3.1.11,**

Figure 1E). We observed a diminished diversity of Ig repertoires in both lymph node and spleen of *Aire*$^{-/-}$ rats (P = 0.041, **Figure 1D**), implying a greater extent of selective expansion during the GC reaction. To confirm this, we also calculated Chao1 and D50 diversities which are two other diversity indices commonly used in immune repertoire research (Chaudhary & Wesemann, 2018). D50 and Chao1 also showed a significant decline in immunoglobulin repertoire diversity (**Figure S4**), providing further evidence of clonal expansions. Because the amount and strength of CD40- and cytokine-mediated signals received by GC B cells from cognate TFH cells defines the magnitude of subsequent B-cell expansion and hypermutation (Gitlin et al., 2014), these results suggest that B cells may be chronically activated in secondary lymphoid organs as a result of increased numbers of cognate autoreactive T cells. This hyperactivation may in turn allow autoreactive B cells to mature and seed the memory and autoantibody-secreting plasma cell compartments.

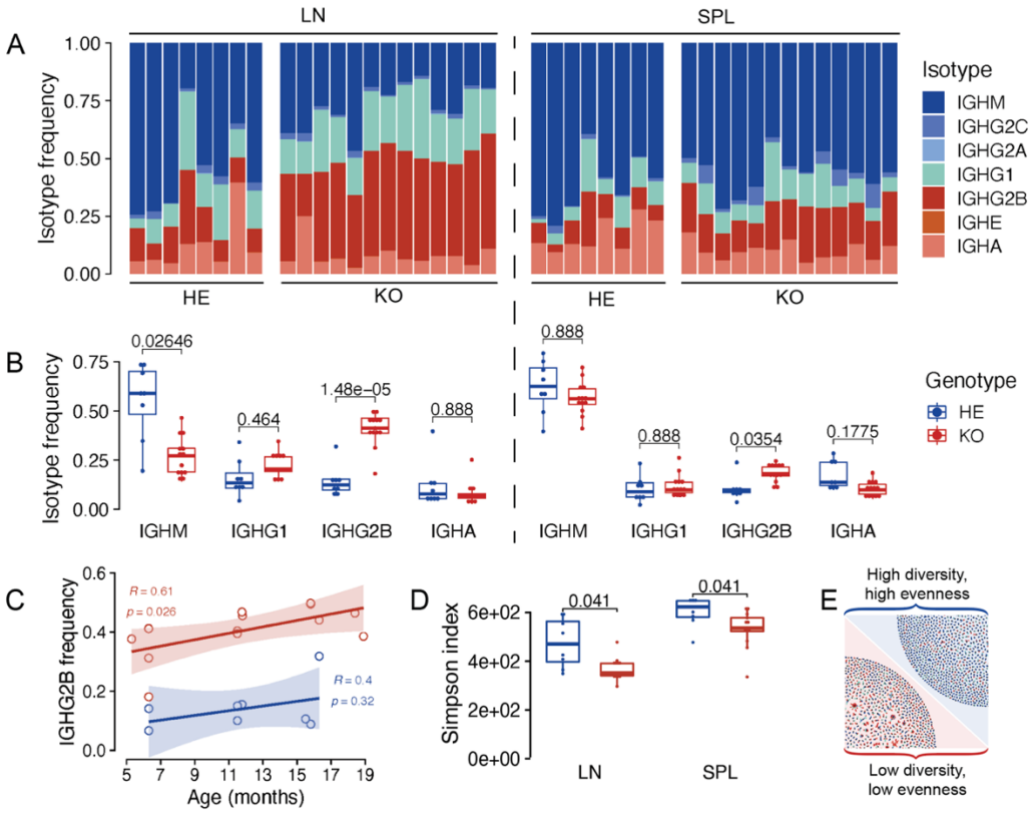

Figure 7. Isotype frequencies and clonality. (A-B) Isotype frequencies in each sample expressed as percentage of unique VDJ sequences per isotype. In the legend, immunoglobulin classes are ordered according to the genomic position of the

corresponding C region. (**C)** IgG2b isotype frequency as a function of the animal age for *Aire*$^{-/-}$ (red) and *Aire*$^{+/-}$ (blue) rats. (**D)** Simpson diversity of immunoglobulin repertoires calculated on clonotype distributions as outlined in **Methods 3.1.11**. (**E)** Explanation of diversity in the form of network representations of Ig repertoires in which each edge vertex represents a single VDJ sequence (clone) and edges join vertices representing ancestor-child pairs that result from somatic hypermutation of class switch recombination. Network topologies were inferred as outlined in **Methods** (section 3.1.10). LN – lymph node; SPL – spleen; HE – *Aire*$^{+/-}$; KO – *Aire* $^{-/-}$. Holm-Bonferroni corrected p values from Student's t-test are shown.

3.2.2 Skewed IGHV gene usage in IgG2b repertoires

B cell selection in GCs is known to be affinity-driven and the amount of antigen captured and presented to T cells during GC reaction has been shown to influence the degree of subsequent B-cell proliferation in the dark zone (Gitlin et al., 2014). CDRs of the Ig variable domain, two of which are encoded by the Ig V gene, are considered to be the most important for antigen recognition, and some Ig V genes have been previously associated with autoimmunity (Bashford-Rogers et al., 2019; Watson & Breden, 2012).

For this reason, we next examined the frequency of Ig V genes in all major Ig subtypes (IgM, IgG1, IgG2b, IgA) separately. We found that the usage of IGHV1-18, IGHV5-25, and IGHV5-27 genes was skewed among IgG2b clones in lymph nodes of *Aire*$^{-/-}$ animals (**Figure 2A**). IGHV1-18 gene rose to a frequency of 1.51% despite being poorly represented in healthy animals, with an average frequency of 0.2% across both immune organs ($P < 0.05$). Thus, IGHV1-18 may represent an autoantigen-associated sequence signature that is kept at a low frequency due to T-cell-dependent peripheral tolerance checkpoints but becomes expanded as a result of T cell tolerance breakdown. IGHV5-25 and IGHV5-27 genes also increased in frequency by approximately 1% each ($P < 0.05$). Interestingly, neither of these two genes were differentially expressed in other isotype groups or when different Ig subtypes where pooled (**Fig S2A**), suggesting that IGHV5-25 and IGHV5-27 clones most commonly switched to IgG2b and were mostly confined to the memory cell compartment. In addition, we observed a high degree of sequence homology between these two genes (**Figure 2**). Within CDR1 and CDR2 regions, only one site was different at position 64, and both residues (serine and asparagine) were polar amino acids (**Figure S3B**). This functional

similarity of IGHV5-25 and IGHV5-27 genes indicates that their co-expansion may be a result of autoantigen-driven convergence in which phylogenetically unrelated clones of the same specificity undergo affinity maturation within GCs.

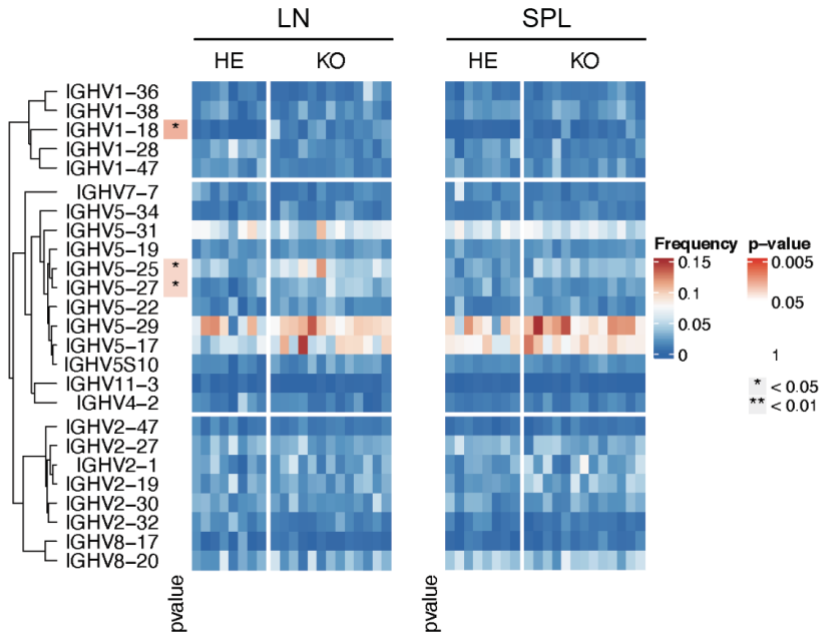

Figure 8. Ig V gene usage in the IgG2b repertoire of *Aire*$^{-/-}$ rats. The frequency of clones expressing a given V gene is shown on the heatmap. V genes were selected by filtering out dropouts (i.e. genes which were not detected in all of the samples) and rare genes whose frequency was less than 1%. Gene statistics were computed at the clonotype level without taking into consideration clone size. Dendrogram on the left illustrates functional similarity between different V genes and was obtained by performing average linkage clustering of V gene amino acid sequences. ** $p \leq 0.01$, * $p \leq 0.05$, two-tailed Student's t-test with Holm Bonferroni correction.

3.2.3 Functional convergence of immunoglobulin repertoires

Next, we sought to examine clonotype sharing between lymph node and splenic Ig repertoires in individual rats. Overlap of immune repertoires between different individuals is often used as a readout of functional convergence, for example in response to immunodominant viral epitopes (Galson et al., 2020). However, within the same organism, repertoire similarity can reflect the extent of cell proliferation as well as lymphocyte

recirculation (Meng et al., 2017a). In rat, B cells can freely recirculate between blood and lymph through a specialized subset of blood vessels called post-capillary venules (James Learmonth Gowans and E. J. Knight, 1964). This homing of B cells into the lymph nodes implies that lymphocyte repertoires in peripheral lymphoid tissues are highly interconnected. For this reason, clonotype sharing is often used as a metric to assess etiological proximity of lymphocyte repertoires within the same individual (Izraelson et al., 2018).

Compared to repertoires of the same animal, we observed only minimal clonotype overlap between healthy littermates, suggesting that each animal has its own "private" lymphocyte repertoire (**Figure 7B**). When compared to healthy controls, *Aire*$^{-/-}$ rats exhibited a roughly two-fold higher degree of overlap between secondary lymphoid organs ($P < 0.05$). We next tried to dissect the composition of the shared repertoire by calculating clonotype overlaps of the N most abundant clones with incrementally growing N. This allowed us to determine the relationship between clonotype expansion status and its likelihood of sharing between secondary lymphoid organs. Interestingly, overlap of Ig repertoires in *Aire*$^{-/-}$ rats was substantially diminished when only top 100 expanded clones were considered, indicating that SLOs have unique, largely nonoverlapping, sets of hyperexpanded clones (**Figure 7B**). This suggests that lymph node and spleen may represent two antigenically distinct compartments. It is possible that lymph- and blood-borne self-proteins which are recycled through the lymphatics and spleen, respectively, represent two distinct subsets of the entire autoantigen repertoire which results in the expansion of B cells with largely non-overlapping specificities. Less frequent clonotypes, on the other hand, likely represent B cells that recirculate between blood and lymph and may reflect each animal's previous immunological history.

3.2.4 Dysregulated isotype switching

Similarly to SHM, CSR also happens in response to T-cell dependent cytokine stimulation, and abnormal isotype switching has been reported in a number of inflammatory diseases including Behçet's disease, systemic lupus erythematosus (SLE), and Crohn's disease. Therefore, we asked whether dysregulated CSR was also characteristic of our APS-1 model and whether it could explain the increased frequency of IgG2b subtype. To test this, we implemented an algorithm which reconstructs class-switch events from Ig repertoire data (**Methods 3.1.10**). Briefly, the algorithm builds a minimum spanning tree from a weighted graph of possible ancestor-child pairs and searches for ancestor-child pairs having different

classes (Horns et al., 2016). For each possible class-switch event we then determined probabilities that define the class switching landscape.

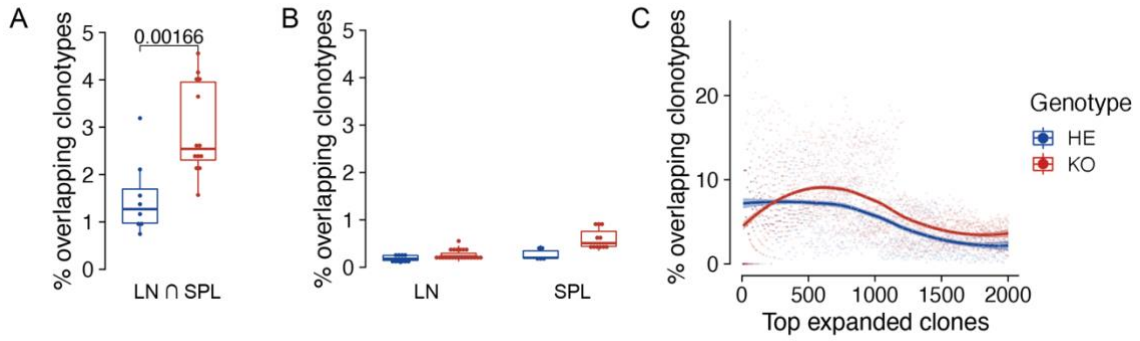

Figure 9. Ig repertoire overlap. (A) Percentage of overlapping clonotypes between samples from two lymphoid tissues of the same animal and **(B)** between Ig repertoires of littermates; Holm-Bonferroni corrected p values from Student's t-test are shown. **(C)** Percentage of overlapping clonotypes as a function of clonal expansion status. Ig repertoires were sampled to include N most expanded clones with incrementally growing N and percentage overlap was calculated at each increment. In order to correct for differential sample coverage, each value in (**A-C)** represents a bootsrap mean of 100 independent resamplings. Clonotypes are defined as shared between samples if each sample contributed at least one sequence to the clonotype. LN – lymph node; SPL – spleen.

In both immune organs, we observed an overall increase in switching from IgM to IgG2b and from IgG1 to IgG2b in lymph nodes of *Aire*$^{-/-}$ animals, which was consistent with increased frequency of IgG2b (**Figure 10A**). We also observed an increased tendency for IgM and IgG1 classes to switch to the IgG2b subtype (**Figure S5**; destination probability; $P < 0.05$). On average, 73% of IgG2b clones in lymph nodes of *Aire*$^{-/-}$ animals had naïve IgM class as their immediate ancestor, implying that switching rather than expansion alone was responsible for the increase in IgG2b frequency. At the same time, there was a significant reduction in switching of IgG2b to downstream classes such as IgA (**Figure S5;** destination probability; $P < 0.05$). On average, about 79% of IgG2b clones remained in the same class in the next generation (compared to 50% in controls), suggesting a stronger antigen-driven selection for those sequences.

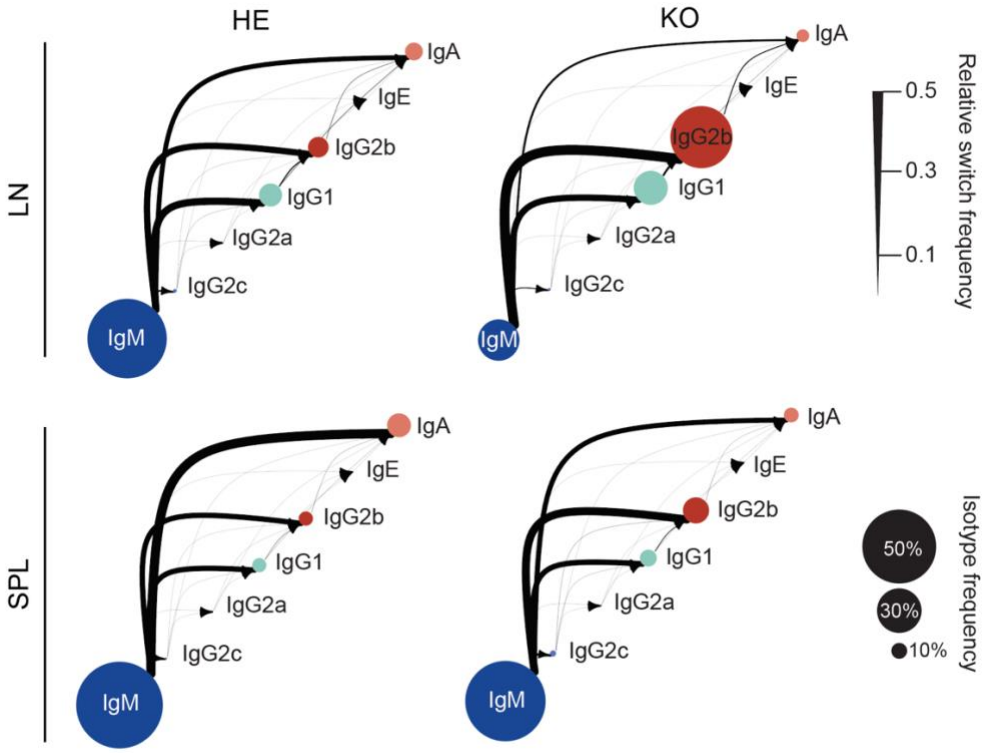

Figure 10. Class switching probabilities. State transition diagram of immunoglobulin class switching in secondary lymphoid tissues of *Aire*$^{-/-}$ rats. Isotypes are indicated as circles and possible switch event as arrows. The order of classes from left to right reflects the organization of respective C regions in the germline. The radius of each circle indicates relative frequency of the labeled class. The width of each arrow indicates the relative frequency of the switch. LN – lymph node; SPL – spleen.

3.2.5 scRNA-seq of splenocytes identifies multiple immune cell types

To understand regulatory and possible pathogenic mechanisms of autoantibodies in the APS1 rat model, we performed single-cell RNA sequencing on splenic CD45^{+} cells which were purified by flow cytometry from 6-months-old *Aire*$^{-/-}$ and *Aire*$^{+/-}$ rats. After quality control and doublet removal, we obtained a total of about 18,400 cells from 3 *Aire*$^{+/-}$ and 3 *Aire*$^{-/-}$ rats (**Figure 11C**). To correct for technical differences between samples and to perform comparative analysis across conditions, we performed batch correction using the Seurat v4 algorithm which identifies correspondences between cell states that are shared across different samples (Stuart et al., 2019).

To infer immune cellular composition in our data, we annotated cell clusters into 18 different immune cell types or cell states using previously described marker gene sets (**Figure 11A-B**). Among myeloid lineage cells, we found macrophages (*Cd86*, *Apoe*, *Adgre1*) which further subclustered into M1-like (*Csfr1*, *Cd14*) and M2-like (*Cd51*, *C1qb*, *Cd169*, *Mrc1*) macrophages, conventional dendritic cells (*Cd11c*, *Itgax*, *Cd86*, *Cd83*), plasmacytoid dendritic cells (*Cd4*, *Siglech*) and neutrophils (*Ptprc*, *S100a8*). Using ImmGen data browser, we identified several B cell subtypes including Transitional type 2 (*Hmgb2*, *Rgs13*) B cells, marginal zone (MZ; *Cd19*, *Cd22*, *Cr2*) B cells, follicular (*Fcer2*) B cells, two memory B cell clusters (*Cd79a*, *Ms4a1*; *Mx1*, *Mx2*), and a cluster of plasma cells (*Tifa*). We found a cluster of NK cells and additionally a cluster of activated NK cells which was defined by high expression of *Xcl1*, *Ctsw*, and *Ifitm1*. T cells were represented by 4 major subtypes including naïve CD4$^+$ T cells (*Sell*), CD4+ T_{EFF} (*Il2ra*, *Foxp3*, *Icos*) cells, CD8$^+$ naïve T cells (*Cd8a*, *Sell*) and CD8$^+$ T_{CM} cells (*Gzmk*, *Ccl5*, *Ctsw*).

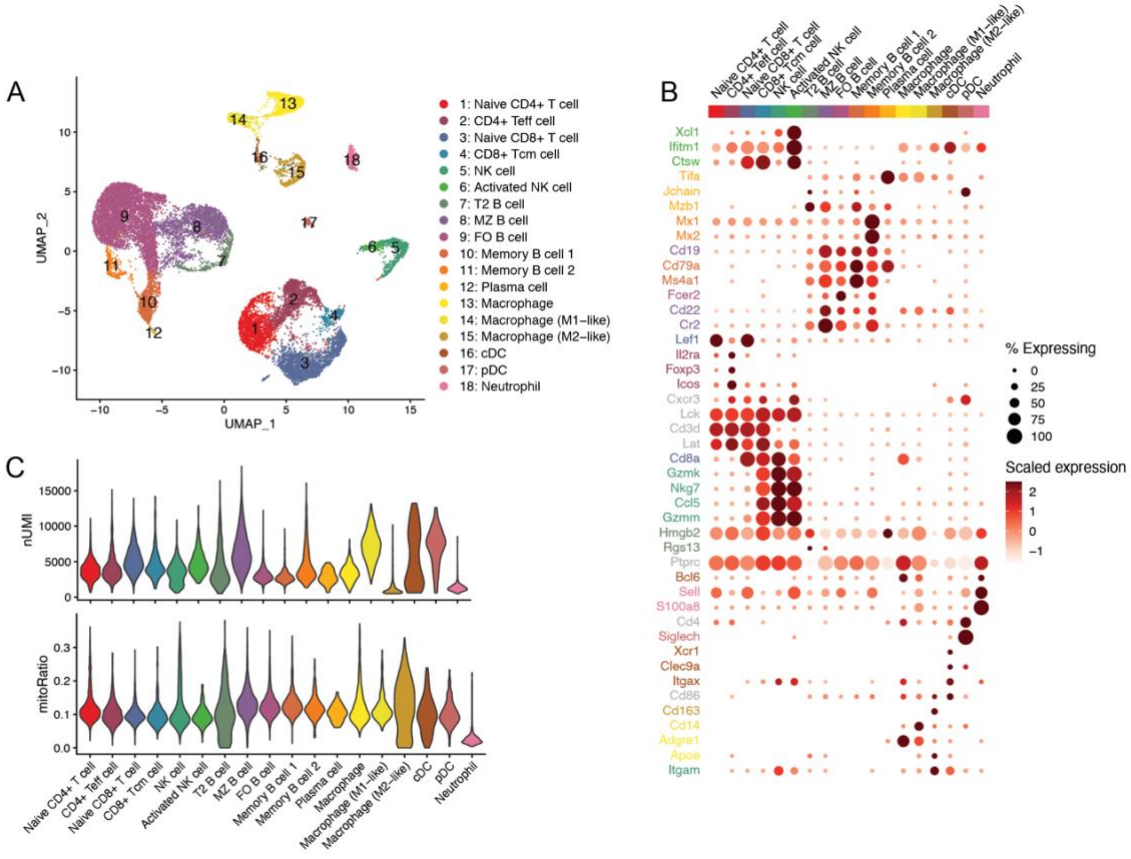

Figure 11. Immune cell heterogeneity of rat spleen. (**A**) UMAP embedding of scRNA-seq data colored by cell type. CD4$^+$ T_{EFF} cell – CD4+ T effector cell; CD8$^+$ T_{CM} cell – CD8$^+$ T central memory cell; NK cell – natural killer cell; T2 B cell – transitional type 2 cell; MZ B cell – marginal zone B cell; FO B cell – follicular B cell; cDC – conventional dendritic

cell; pDC – plasmacytoid dendritic cell. (**B**) Dot plot for expression of marker genes in rat spleen. Color represents scaled normalized mean expression of marker genes in each cell type, and size indicates the fraction of cells expressing marker genes. (**C**) Violin plots showing the distribution of unique molecular identifiers (UMIs) detected in each cell and the proportion of transcripts that map to the mitochondrial genome. Apoptotic or stressed cells often exhibit extensive mitochondrial contamination.

3.2.6 Type I interferon gene signature is diminished in AIRE-deficient rats

We next compared cell transcriptomes between *Aire*$^{-/-}$ and *Aire*$^{+/-}$ rats by performing pseudobulk differential expression testing using *edgeR* (Robinson et al., 2009). Altogether, we found 142 genes which were differentially expressed across 16 cells types (**Figure 12A**). Gene ontology analysis revealed enrichment of pathways related to defense response (GO:0051607, p = 1.34e-14) and response to interferon alpha (GO:0035455, p = 1.18e-09), suggesting an impaired IFN-α signaling pathway (**Figure 12B**). To confirm the presence of a diminished IFN transcriptional signature, we analyzed the expression of interferon stimulated genes (ISGs). We measured the intensity of IFN signature by computing "IFN score" which integrates normalized relative expression of the core ISG set (**Methods 3.1.13**), as previously described (Nezos et al., 2015). We indeed observed a diminished IFN signature in all of the cell types, especially among neutrophils and DCs which has been reported previously (Mostafavi et al., 2016), suggesting a biased and cell-type specific effect (**Figure 12C**).

We hypothesized that this effect was due to IFN-α neutralization by anti-IFN-α autoantibodies which we previously found to be present in *Aire*$^{-/-}$ rats of this age (**Figure 6C**). The presence of strongly neutralizing anti-IFN autoantibodies would in turn result in the loss of tonic IFN signaling and under-expression of ISGs that are sensitive to tonic IFN signaling at baseline (i.e. under homeostatic conditions in pathogen-free rats). Indeed, when we analyzed the expression of tonic-sensitive ISGs which were identified previously in *Ifnar1*-deficient mice (Mostafavi et al., 2016), we found that ISGs most sensitive to tonic IFN were also most strongly downregulated in B cells and macrophages of AIRE-deficient rats (**Figure 13A**). In addition, we analyzed transcriptomes of bulk CD4^{+} T cell subtypes using microarray technology and found concordant results, suggesting that anti-IFN autoantibodies effectively neutralize IFN-α and act on cells by diminishing the expression

of tonic-sensitive ISGs (**Figure 13B**). Moreover, GO analysis of microarray data revealed subcategories related to defense response and regulation of interferon production (**Figure 13C**), providing evidence that T cells are similarly affected by anti-IFN autoantibodies.

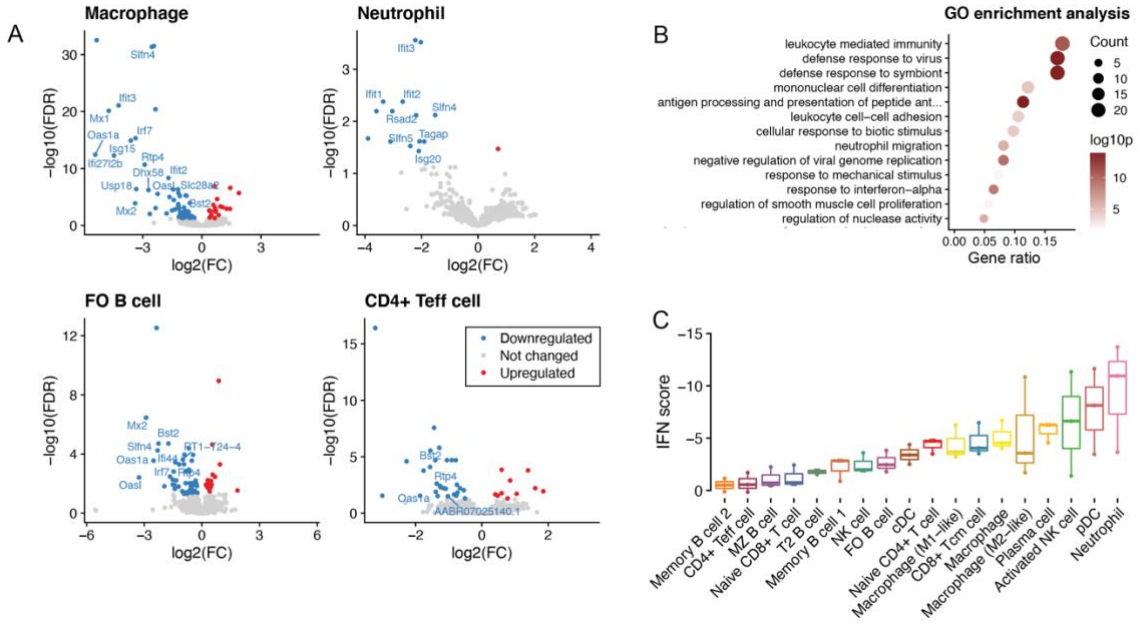

Figure 12. Interferon gene signatures in immune cell types. (**A**) Volcano plots comparing pseudobulk gene expression of *Aire*$^{-/-}$ and *Aire*$^{+/-}$ rats in several immune cell types. Interferon-stimulated genes are labeled (FO B cell, follicular B cell; CD4$^+$ TEFF cell, CD4$^+$ T effector cell). (**B**) Gene ontology (GO) enrichment analysis of genes differentially expressed across all cell types. Only 13 most enriched GO terms are shown. Gene ratio corresponds to the percentage of all differentially expressed genes in the given GO term. Hypergeometric test p values are shown. (**C**) Type I interferon (IFN) signature score calculated on relative expression of a set of interferon-stimulated genes (ISGs; **Methods 3.1.13**). Negative IFN score indicates downregulation of ISGs and a diminished IFN transcriptional signature. CD4$^+$ TEFF cell – CD4+ T effector cell; CD8$^+$ TCM cell – CD8$^+$ T central memory cell; NK cell – natural killer cell; T2 B cell – transitional type 2 cell; MZ B cell – marginal zone B cell; FO B cell – follicular B cell; cDC – conventional dendritic cell; pDC – plasmacytoid dendritic cell.

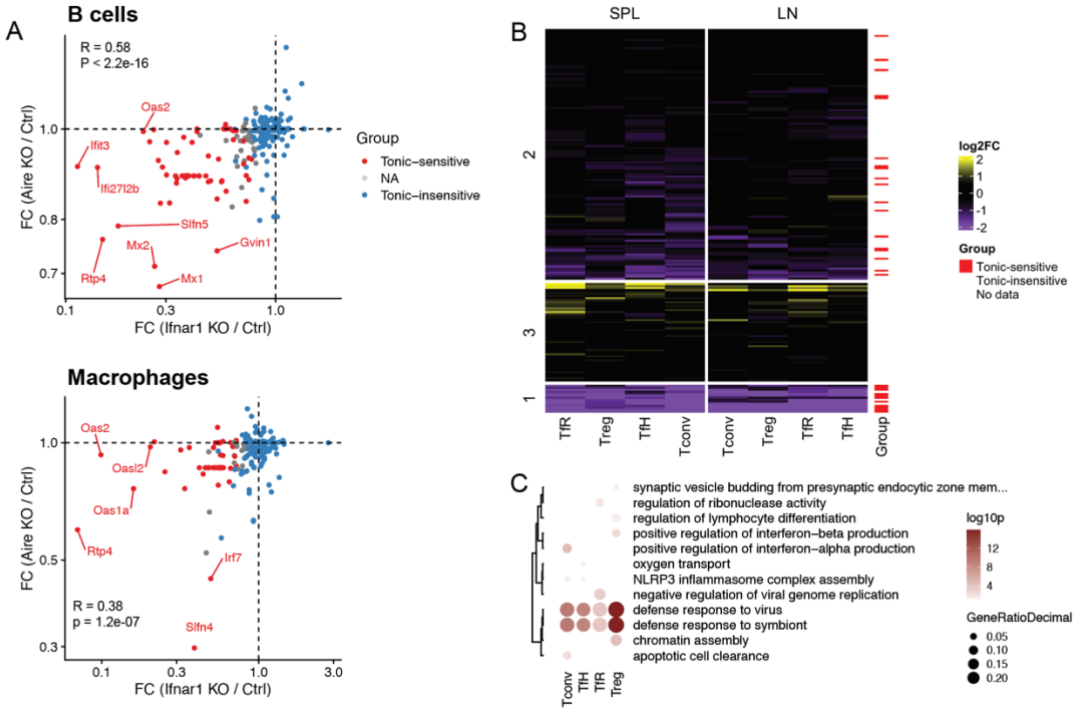

Figure 13. Tonic-sensitivity of differentially expressed interferon-stimulated genes. (**A**) Scatterplots comparing relative gene expression of interferon-stimulated genes (ISGs) in B cells and macrophages in *Aire*$^{-/-}$ rats and *Ifnar1*$^{-/-}$ mice. Tonic sensitivity of ISGs was inferred in the publication by (Mostafavi et al., 2016). Correlation coefficient R and associated p value are shown in both figures. (**B**) Heatmap showing relative expression of ISGs in CD4^{+} T cell population in lymphoid tissues of AIRE-deficient rats. The bar on the right indicates tonic sensitivity of an ISG (SPL, spleen; LN; lymph node). (**C**) Gene ontology (GO) enrichment analysis of differentially expressed genes CD4^{+} T cell populations in *Aire*$^{-/-}$ rats. Gene ratio corresponds to the percentage of all differentially expressed genes in the given GO term. Hypergeometric test p values are shown with color. Dendogram on the left represents semantic similarity of GO terms which was determined using *GOSemSim* package (Yu et al., 2010). T_{conv} – T conventional cells; T_{reg} – T regulatory cells; T_{FH} – T follicular helper cells; T_{FR} – T follicular regulatory cells.

3.2.7 NK cell effector function is impaired in AIRE-deficient rats

Next, we analyzed compositional changes in our scRNA-seq data using Milo R package which uses a k-nearest neighbors (KNN) graph to define groups of connected cells (hereafter referred to as neighborhoods) and tests for differential abundance in cell neighborhoods

(Dann et al., 2022). This has an advantage over using discrete clusters as input, since cell clusters often fail to provide the appropriate resolution and cannot capture continuous trajectories. Milo identified 717 neighborhoods spanning the KNN graph, of which 3 showed evidence of differential abundance (FDR < 0.05; **Figure 14A**). All three neighborhoods mapped to the cluster containing activated NK cells, suggesting that NK cell activation was impaired. To identify gene expression signatures associated with this change, we further subclustered NK cells into activated (*Xcl1*, *Ctsw*) and resting NK cells, defined by low expression of ribosomal proteins (**Figure 14B**), and inferred pseudotime trajectories describing NK cell activation and inhibition (**Figure 14C**, **Methods 3.1.13**). We then examined genes that correlated with activation trajectory and found ISGs (*Ifi27lb*, *Ifitm2*) that were co-expressed with genes related to NK cell effector function, such as *Xcl1*, *Cd27*, *Sell*, and *Ctsw* (Hayakawa & Smyth, 2006; Kroczek et al., 2018; Stoeckle et al., 2009), suggesting a possible role of type I interferons in NK cell activation (**Figure 14D**).

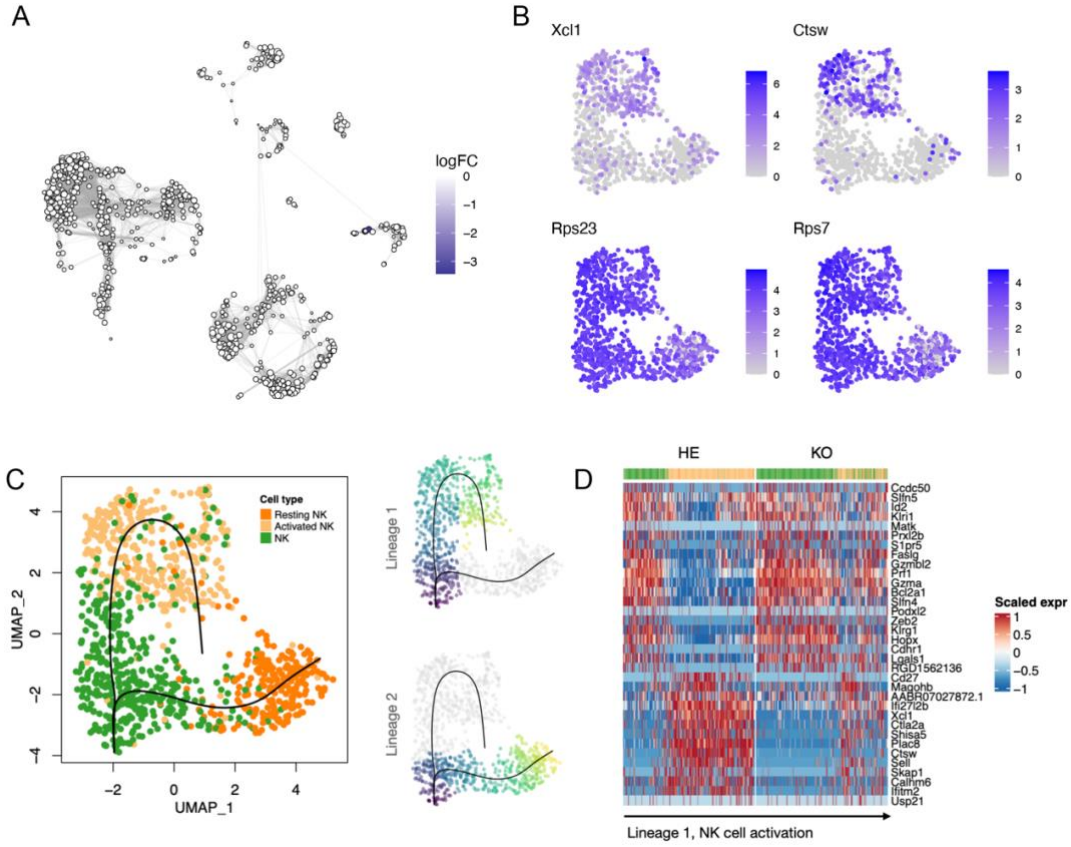

Figure 14. Compositional and associated transcriptional changes in NK cells of AIRE-deficient rats. (**A**) An abstracted graph of neighborhoods superimposed on the UMAP embedding of the data. Color represents log fold change (logFC) estimated with differential

neighborhood abundance testing, node size represents the size of the neighborhood, and edge width represents the amount of overlap between neighboring neighborhoods. Relative changes which were not statistically significant ($p > 0.05$) are not shown. (**B**) Expression of marker genes for NK cell subsets. (**C**) Left: a UMAP embedding of NK cells colored by cell state together with principal curves of each lineage. Right: UMAP embeddings of NK cells colored by the inferred position in the pseudotemporal ordering. (**D**) Heatmap showing scaled normalized expression of genes associated with pseudotime. Rows represent genes and columns represent single cells, ordered by pseudotime from left to right. The bar on the top shows the distribution of cell types colored by cell states, same as in (**C**).

3.3 DISCUSSION

In this work, we applied Ig repertoire sequencing and scRNA-seq to study autoantibody repertoires and their possible immunomodulatory role. Ig repertoire analysis in particular has revealed a number of disease-specific signature patterns of antibody responses, including significant clonal expansions and biased responses to IgG2b subtype.

Subdivision of IgG antibodies into several subtypes has been recognized in multiple species, and there exist differences in the ability of different subtypes to exert Fc-dependent functions such as complement activation, antibody-dependent cytotoxicity (ADCC), and opsonization. In the past, rat IgG2b monoclonal antibodies have received a lot of interest for immunotherapy due to their ability to fix human complement and promote ADCC with human effector cells (Hale et al., 1985; Medgyesi et al., 1980). For example, rat IgG2b anti-CD52 mAbs (CAMPATH-1 G) have been shown to produce long-lasting depletion of lymphocytes in patients with lymphoid malignancies (Dyer et al., 1989). In this regard, rat IgG2b behaves like human IgG1 and IgG3 in binding to Fc-gamma receptor FcγRI and mediating ADCC through FcγRIII expressed on NK cells (Hale et al., 1985). Overall, IgG2b antibodies clearly possess very potent effector functions in concert with human immune cells. Although data on *in vivo* activity of IgG2b antibodies in rats is lacking, it is possible to speculate that in the context of AIRE deficiency, IgG2b response against self-antigens is accompanied by tissue-destructive ADCC reactions which leads to the severe autoimmune phenotype observed in AIRE-deficient animals.

Generally, the class of antibody produced by a B cell depends on the T helper subset involved, and it has been shown that IgG2b production is mediated by T_H1 cells that produce

IL-2 and IFN-γ (Gracie & Bradley, 1996; Saoudi et al., 1993). Conversely, the production of IL-4 by T_H2 cells causes generation of IgG1 and IgE antibodies (Guéry Bastien Cautain et al., 1999). Whereas T_H2 cells are thought to be protective, polarized T_H1 responses have been implicated in the pathogenesis of a number of model autoimmune diseases such as experimental autoimmune encephalomyelitis (EAE) and insulin-dependent diabetes mellitus (IDDM) (Liblau et al., n.d.). Importantly, $CD4^+$ effector/memory T cells of APS1 patients have been found to display abnormalities in the IFN-γ pathway which is itself a canonical T_H1 cytokine (Heikkilä et al., 2016). Thus, it is possible that generation of IgG2b antibody subclass in *Aire*$^{-/-}$ rats represents a signature of a T_H1-dominated immune response caused by the loss of central T cell tolerance. T_H1/T_H2 polarization is influenced by the cytokine milieu at the time of the antigenic challenge, among other factors (Rutz & Ouyang, 2011). This implies that the manipulation of T_H1/T_H2 balance by therapeutically changing the cytokine profile of autoreactive $CD4^+$ T cells may be a feasible immunotherapeutic approach to ameliorate some of the autoimmune symptoms in AIRE-deficient rats.

These findings are in contrast to another study by Ossart et al. (2018) which identified increased titers of IgM, IgG1, IgA, and not IgG2b antibodies in sera of *Aire*$^{-/-}$ rats. This discrepancy may be related to ages of animals, their genetic background and animal facilities in which they are kept. Importantly, peripheral blood and secondary lymphoid organs represent different anatomical compartments which may contain distinct Ig repertoires. This tissue-specificity of Ig repertoires has been shown in a study by Meng et al. (2017) who found that in blood-rich tissues each branch of the B-cell network was tissue-specific. It is also possible that by analyzing Ig repertoires from lymphoid organs we are detecting purer expansion signals from plasma and activated memory B cells (Yang et al., n.d.).

Although our experimental design did not allow us to infer antigen-binding specificities of IgG2b antibodies, it is possible to speculate that at least some of IgG2b clones may be autoreactive. This is evidenced by significant biases in V gene segment usage that we detected among antigen-specific IgG2b clonotypes, suggesting that these responses may have been directed against certain immunodominant epitopes of *Aire*$^{-/-}$ rat autoantigens. To provide a better understanding of the role of IgG2b clones in the pathogenesis of APS1-like symptoms in AIRE-deficient rats, further studies may study these cells with more sophisticated approaches capable of mapping of Ig sequence to antigen specificity (Setliff et al., 2019).

Using scRNA-seq, we found that anti-IFN autoantibodies act on cells by effectively neutralizing Type I IFNs and diminish the expression of ISGs. These findings are consistent with another study that found down-regulation of ISGs in APS1 sera with autoantibodies to IFN-α (Kisand et al., 2008). Type I IFNs such as IFN-α are produced in large quantities by pDCs during acute viral infections which induces the expression of downstream antiviral effectors (Ali et al., 2019). Under homeostatic conditions, however, there is always an intrinsic IFN production in response to the natural shedding of microbial molecules by gut microbiota, which serves to enhance the sensitivity of cells to acutely produced IFN (Abt et al., 2012). These two IFN signaling modes result in distinct transcriptional outcomes in which tonic IFN signals regulate the expression of a distinct subgroup of ISGs with high sensitivity to low IFN doses but with slower response kinetics (Mostafavi et al., 2016). In this work, we show that deprivation of basal levels of circulating IFNs by the neutralizing anti-IFN-α autoantibodies mostly affects tonic-sensitive ISGs, raising a possibility that AIRE-deficient individuals may be more susceptible to microbial infections. However, despite decreased expression of ISGs in blood cells, viral infections in APS1 patients are surprisingly rare, suggesting a possible compensatory mechanism by IFNs other than IFN-α (Kisand et al., 2008).

Type I interferons are known to be an important link between innate and adaptive immunity by acting on several lymphoid cell subsets. For example, Type I IFNs are known to stimulate NK cell proliferation, their cytotoxicity and cytokine secretion and influence adaptive immune responses indirectly by upregulating MHC class I and class II expression by dendritic cells (Madera et al., 2016). Our own results suggest that in the context of AIRE deficiency, inhibition of ISG responses by neutralizing anti-IFN-α autoantibodies impairs activation of NK cells which may lead to a broader immunosuppressive effect.

Such disease-ameliorating properties of APS1 autoantibodies have been proposed in the past, albeit via a different mechanism. Type I IFNs are known to contribute to the onset of type 1 diabetes (T1D), and in one of the previous studies by our lab group it has been noted that patients with strongly neutralizing anti-IFN-α antibodies failed to develop T1D (Meyer et al., 2016). Results presented in this work suggest that immunosuppression via impaired activation of NK cells, which are known to be important mediators in several immunopathologies (Segerberg et al., 2019), may be another mechanism of direct disease amelioration by anti-IFN-α autoantibodies.

For many years research into APS1 has been hampered due to the scarcity of patient samples and the lack of animal models capable of faithfully recapitulating the disease. The recently developed *Aire*$^{-/-}$ rat model, however, more accurately reproduces key phenotypic features of human subjects, which allowed us to design a well-replicated experiment with a greater statistical power. One drawback of such an approach is that it is difficult to assess the extent to which these findings translate to pathophysiology of human APS1 subjects, since some immune system components are known to be species-specific. For example, human Ig locus encodes C regions for four classes of IgG (IgG1, IgG2, IgG3, and IgG4) which have no direct homologs in rats and mice (Mestas & Hughes, 2004). Additionally, many other immune system components are known to be species-specific, such as cytokines, B cell and T cell signaling pathways (Mestas & Hughes, 2004), further complicating the translation of animal models to human studies. Nevertheless, findings presented in this thesis may be useful for hypothesis generation and guiding further studies into APS1 that may choose to expand on this work – for example, by focusing on the role of a closely related human IgG3 antibody subclass and how it correlates with T_H1 and T_H2 responses.

SUMMARY

In this study, we applied immunoglobulin repertoire profiling and single-cell RNA sequencing to characterize antibody responses in AIRE-deficient rats and identify possible immunomodulatory effects of their autoantibody repertoire. We detected strongly biased responses to IgG2b subtype which we propose may represent a signature of a T_H1-dominated immune response caused by the breakdown of T cell tolerance. T_H1 hyperfunction may lead to excessive activation of B cells, generation of autoantibodies, and tissue damage. These findings call for more mechanistic studies examining the role of T_H1/T_H2 imbalance in generation of IgG2b antibodies or their human counterparts in the context of AIRE deficiency, for which the AIRE-deficient rat model will be an invaluable tool.

scRNA-seq of rat splenocytes revealed a pervasive loss of ISG expression due to anti-IFN autoantibodies and allowed us to recover its cell-type-specific effect. Although the diminished numbers of activated NK cells should not come as a surprise, given our knowledge of direct Type I IFN effects on NK cell memory and cytotoxicity (Müller et al., 2017), this finding went largely unnoticed in previous studies due to low resolution and throughput of standard single-cell profiling techniques such as flow cytometry. This finding may have wide-ranging implications for a number of other diseases for which anti-IFN antibodies are known to be pathognomonic. This necessitates further studies to identify their possible immunomodulatory effect on various lymphoid cell subsets in APS1 and other immune-mediated diseases.

Altogether, this analysis of immunoglobulin repertoires and single-cell profiles in AIRE-deficient rats provides important clues about possible etiology, nature, and effects of their autoantibodies and lays groundwork for better understanding of pathological mechanisms in human patients.

REFERENCES

Abt, M. C., Osborne, L. C., Monticelli, L. A., Doering, T. A., Alenghat, T., Sonnenberg, G. F., Paley, M. A., Antenus, M., Williams, K. L., Erikson, J., Wherry, E. J., & Artis, D. (2012). Commensal Bacteria Calibrate the Activation Threshold of Innate Antiviral Immunity. *Immunity*, *37*(1), 158–170. https://doi.org/10.1016/j.immuni.2012.04.011

Ali, S., Mann-Nüttel, R., Schulze, A., Richter, L., Alferink, J., & Scheu, S. (2019). Sources of type I interferons in infectious immunity: Plasmacytoid dendritic cells not always in the driver's seat. In *Frontiers in Immunology* (Vol. 10, Issue APR). Frontiers Media S.A. https://doi.org/10.3389/fimmu.2019.00778

Anderson, M. S., Venanzi, E. S., Klein, L., Chen, Z., Berzins, S. P., Turley, S. J., von Boehmer, H., Bronson, R., Dierich, A., Benoist, C., & Mathis, D. (2002). Projection of an immunological self shadow within the thymus by the aire protein. *Science*, *298*(5597), 1395–1401. https://doi.org/10.1126/science.1075958

Bashford-Rogers, R. J. M., Bergamaschi, L., McKinney, E. F., Pombal, D. C., Mescia, F., Lee, J. C., Thomas, D. C., Flint, S. M., Kellam, P., Jayne, D. R. W., Lyons, P. A., & Smith, K. G. C. (2019). Analysis of the B cell receptor repertoire in six immune-mediated diseases. *Nature*, *574*(7776), 122–126. https://doi.org/10.1038/s41586-019-1595-3

Blanco, E., Pérez-Andrés, M., Arriba-Méndez, S., Contreras-Sanfeliciano, T., Criado, I., Pelak, O., Serra-Caetano, A., Romero, A., Puig, N., Remesal, A., Torres Canizales, J., López-Granados, E., Kalina, T., Sousa, A. E., van Zelm, M., van der Burg, M., van Dongen, J. J. M., & Orfao, A. (2018). Age-associated distribution of normal B-cell and plasma cell subsets in peripheral blood. *Journal of Allergy and Clinical Immunology*, *141*(6), 2208-2219.e16. https://doi.org/10.1016/j.jaci.2018.02.017

Briney, B. S., Willis, J. R., Finn, J. A., McKinney, B. A., & Crowe, J. E. (2014). Tissue-specific expressed antibody variable gene repertoires. *PLoS ONE*, *9*(6). https://doi.org/10.1371/journal.pone.0100839

Chan, T. D., Gatto, D., Wood, K., Camidge, T., Basten, A., & Brink, R. (2009). Antigen Affinity Controls Rapid T-Dependent Antibody Production by Driving the Expansion Rather than the Differentiation or Extrafollicular Migration of Early Plasmablasts. *The Journal of Immunology*, *183*(5), 3139–3149. https://doi.org/10.4049/jimmunol.0901690

Chao, A. (1984). Nonparametric Estimation of the Number of Classes in a Population. In *Scandinavian Journal of Statistics* (Vol. 11, Issue 4).

Chaudhary, N., & Wesemann, D. R. (2018). Analyzing immunoglobulin repertoires. In *Frontiers in Immunology* (Vol. 9, Issue MAR). Frontiers Media S.A. https://doi.org/10.3389/fimmu.2018.00462

Coffey, F., Alabyev, B., & Manser, T. (2009). Initial Clonal Expansion of Germinal Center B Cells Takes Place at the Perimeter of Follicles. *Immunity*, *30*(4), 599–609. https://doi.org/10.1016/j.immuni.2009.01.011

Collins, J. T., & Dunnick, W. A. (1993). Germline transcripts of the murine immunoglobulin γ2a gene: structure and induction by IFN-γ. *International Immunology*, *5*(8), 885–891. https://doi.org/10.1093/intimm/5.8.885

Dann, E., Henderson, N. C., Teichmann, S. A., Morgan, M. D., & Marioni, J. C. (2022). Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nature Biotechnology*, *40*(2), 245–253. https://doi.org/10.1038/s41587-021-01033-z

de Silva, N. S., & Klein, U. (2015). Dynamics of B cells in germinal centres. *Nature Reviews Immunology*, *15*(3), 137–148. https://doi.org/10.1038/nri3804

di Zenzo, G., di Lullo, G., Corti, D., Calabresi, V., Sinistro, A., Vanzetta, F., Didona, B., Cianchini, G., Hertl, M., Eming, R., Amagai, M., Ohyama, B., Hashimoto, T., Sloostra, J., Sallusto, F., Zambruno, G., & Lanzavecchia, A. (2012). Pemphigus autoantibodies generated through somatic mutations target the desmoglein-3 cis-interface. *Journal of Clinical Investigation*, *122*(10), 3781–3790. https://doi.org/10.1172/JCI64413

Dyer, M. J. S., Hale, G., Hayhoe, F. G. J., & Waldmann, H. (1989). Effects of CAMPATH-1 Antibodies In Vivo in Patients With Lymphoid Malignancies: Influence of Antibody Isotype. In *Blood* (Vol. 73, Issue 6).

Fishman, D., Kisand, K., Hertel, C., Rothe, M., Remm, A., Pihlap, M., Adler, P., Vilo, J., Peet, A., Meloni, A., Podkrajsek, K. T., Battelino, T., Bruserud, Ø., Wolff, A. S. B., Husebye, E. S., Kluger, N., Krohn, K., Ranki, A., Peterson, H., … Peterson, P. (2017). Autoantibody repertoire in APECED patients targets two distinct subgroups of proteins. *Frontiers in Immunology*, *8*(AUG). https://doi.org/10.3389/fimmu.2017.00976

Flajnik, M. F., & Kasahara, M. (2010). Origin and evolution of the adaptive immune system: Genetic events and selective pressures. In *Nature Reviews Genetics* (Vol. 11, Issue 1, pp. 47–59). https://doi.org/10.1038/nrg2703

Galson, J. D., Schaetzle, S., Bashford-Rogers, R. J. M., Raybould, M. I. J., Kovaltsuk, A., Kilpatrick, G. J., Minter, R., Finch, D. K., Dias, J., James, L. K., Thomas, G., Lee, W. Y. J., Betley, J., Cavlan, O., Leech, A., Deane, C. M., Seoane, J., Caldas, C., Pennington, D. J., … Osbourn, J. (2020). Deep Sequencing of B Cell Receptor Repertoires From COVID-19 Patients Reveals Strong Convergent Immune Signatures. *Frontiers in Immunology*, *11*. https://doi.org/10.3389/fimmu.2020.605170

Gavanescu, I., Kessler, B., Ploegh, H., Benoist, C., & Mathis, D. (2007). *Loss of Aire-dependent thymic expression of a peripheral tissue antigen renders it a target of autoimmunity*. www.pnas.orgcgidoi10.1073pnas.0700259104

Germain, P. L., Sonrel, A., & Robinson, M. D. (2020). PipeComp, a general framework for the evaluation of computational pipelines, reveals performant single cell RNA-seq preprocessing tools. *Genome Biology*, *21*(1). https://doi.org/10.1186/s13059-020-02136-7

Gitlin, A. D., Shulman, Z., & Nussenzweig, M. C. (2014). Clonal selection in the germinal centre by regulated proliferation and hypermutation. *Nature*, *509*(7502), 637–640. https://doi.org/10.1038/nature13300

Gracie, J. A., & Bradley, J. A. (1996). IL-12-induced IgG subclass switching 1217 Interleukin42 induces interferon-y-dependent switching of IgG alloantibody subclass. In *Eur. J. Immunol* (Vol. 26).

Guéry Bastien Cautain, J.-C., Martinez, K., Druet, P., de Abdelhadi Saoudi, M., Bernard, I., & Hoedemaekers, A. (1999). *Experimental Autoimmune Myasthenia Gravis May Occur in the Context of a Polarized Th1- or Th2-Type Immune Response in Rats*. http://www.jimmunol.org/content/162/12/7189http://www.jimmunol.org/content/162/12/7189.full#ref-list-1

Hafemeister, C., & Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biology*, *20*(1). https://doi.org/10.1186/s13059-019-1874-1

Hale, G., Clark, M., & Waldmann, H. (1985). *with human lymphocytes. antibody-dependent cell-mediated cytotoxicity antibodies: isotype specificity of Therapeutic potential of rat monoclonal*. http://www.jimmunol.org/content/134/5/3056

Hayakawa, Y., & Smyth, M. J. (2006). CD27 Dissects Mature NK Cells into Two Subsets with Distinct Responsiveness and Migratory Capacity. *The Journal of Immunology*, *176*(3), 1517–1524. https://doi.org/10.4049/jimmunol.176.3.1517

Heikkilä, N., Laakso, S. M., Mannerström, H., Kekäläinen, E., Saavalainen, P., Jarva, H., & Arstila, T. P. (2016). Expanded CD4+ Effector/Memory T Cell Subset in APECED Produces Predominantly Interferon Gamma. *Journal of Clinical Immunology*, *36*(6), 555–563. https://doi.org/10.1007/s10875-016-0302-5

Horns, F., Vollmers, C., Croote, D., Mackey, S. F., Swan, G. E., Dekker, C. L., Davis, M. M., & Quake, S. R. (2016). Lineage tracing of human B cells reveals the in vivo landscape of human antibody class switching. *ELife*, *5*(AUGUST). https://doi.org/10.7554/eLife.16578.001

Izraelson, M., Nakonechnaya, T. O., Moltedo, B., Egorov, E. S., Kasatskaya, S. A., Putintseva, E. v., Mamedov, I. Z., Staroverov, D. B., Shemiakina, I. I., Zakharova, M. Y., Davydov, A. N., Bolotin, D. A., Shugay, M., Chudakov, D. M., Rudensky, A. Y., & Britanova, O. v. (2018). Comparative analysis of murine T-cell receptor repertoires. In *Immunology* (Vol. 153, Issue 2, pp. 133–144). Blackwell Publishing Ltd. https://doi.org/10.1111/imm.12857

Kadota, Y., Okumura, M., Miyoshi, S., Kitagawa-Sakakida, S., Inoue, M., Shiono, H., & Maeda[2], Y. (2000). Altered T cell development in human thymoma is related to impairment of MHC class II transactivator expression induced by interferon-gamma (IFN-g). In *Blackwell Science* (Vol. 59). https://academic.oup.com/cei/article/121/1/59/6461635

Kisand, K., Bøe Wolff, A. S., Podkrajšek, K. T., Tserel, L., Link, M., Kisand, K. v., Ersvaer, E., Perheentupa, J., Erichsen, M. M., Bratanic, N., Meloni, A., Cetani, F., Perniola, R., Ergun-Longmire, B., Maclaren, N., Krohn, K. J. E., Pura, M., Schalke, B., Ströbel, P., … Meager, A. (2010). Chronic mucocutaneous candidiasis in APECED or thymoma patients correlates with autoimmunity to Th17-associated cytokines. *Journal of Experimental Medicine*, *207*(2), 299–308. https://doi.org/10.1084/jem.20091669

Kisand, K., Link, M., Wolff, A. S. B., Meager, A., Tserel, L., Org, T., Murumägi, A., Uibo, R., Willcox, N., Podkrajšek, K. T., Battelino, T., Lobell, A., Kämpe, O., Lima, K., Meloni, A., Ergun-Longmire, B., Maclaren, N. K., Perheentupa, J., Krohn, K. J. E., … Peterson, P. (2008). Interferon autoantibodies associated with AIRE deficiency decrease the expression of IFN-stimulated genes. *Blood*, *112*(7), 2657–2666. https://doi.org/10.1182/blood-2008-03-144634

Kroczek, A. L., Hartung, E., Gurka, S., Becker, M., Reeg, N., Mages, H. W., Voigt, S., Freund, C., & Kroczek, R. A. (2018). Structure-Function Relationship of XCL1 Used for in vivo Targeting of Antigen Into XCR1+ Dendritic Cells. *Frontiers in Immunology*, *9*. https://doi.org/10.3389/fimmu.2018.02806

Lefranc, M. P., Giudicelli, V., Kaas, Q., Duprat, E., Jabado-Michaloud, J., Scaviner, D., Ginestoux, C., Clément, O., Chaume, D., & Lefranc, G. (2005). IMGT, the international ImMunoGeneTics information system®. *Nucleic Acids Research*, *33*(DATABASE ISS.). https://doi.org/10.1093/nar/gki065

Liblau, R. S., Singer, S. M., & Mcdevitt, H. 0. (n.d.). *Thl and Th2 CD4+ T cells in the pathogenesis of organ-specific autoimmune diseases*.

Madera, S., Rapp, M., Firth, M. A., Beilke, J. N., Lanier, L. L., & Sun, J. C. (2016). Type I IFN promotes NK cell expansion during viral infection by protecting NK cells against fratricide. *Journal of Experimental Medicine*, *213*(2), 225–233. https://doi.org/10.1084/jem.20150712

Meager, A., Peterson, P., & Willcox, N. (2008). Hypothetical review: Thymic aberrations and type-I interferons; attempts to deduce autoimmunizing mechanisms from unexpected clues in monogenic and paraneoplastic syndromes. In *Clinical and Experimental Immunology* (Vol. 154, Issue 1, pp. 141–151). https://doi.org/10.1111/j.1365-2249.2008.03739.x

Meager, A., Wadhwa, M., Dilger, P., Bird, C., Thorpe, R., Newsom-Davis, J., & Willcox, N. (2003). Anti-cytokine autoantibodies in autoimmunity: preponderance of neutralizing autoantibodies against interferon-alpha, interferon-omega and interleukin-12 in patients with thymoma and/or myasthenia gravis. In *Clin Exp Immunol* (Vol. 132). https://academic.oup.com/cei/article/132/1/128/6460674

Medgyesi, G. A., Bazin, H., & Gergely, J. (1980). DIFFERENCES IN THE ABILITY OF RAT IgG SUBCLASSES TO CONSUME COMPLEMENT IN HOMOLOGOUS AND HETEROLOGOUS SERUM. In *Immunology Letters* (Vol. 1).

Meffre, E., & Wardemann, H. (2008). B-cell tolerance checkpoints in health and autoimmunity. In *Current Opinion in Immunology* (Vol. 20, Issue 6, pp. 632–638). https://doi.org/10.1016/j.coi.2008.09.001

Meng, W., Zhang, B., Schwartz, G. W., Rosenfeld, A. M., Ren, D., Thome, J. J. C., Carpenter, D. J., Matsuoka, N., Lerner, H., Friedman, A. L., Granot, T., Farber, D. L., Shlomchik, M. J., Hershberg, U., & Luning Prak, E. T. (2017a). An atlas of B-cell clonal distribution in the human body. *Nature Biotechnology*, *35*(9), 879–886. https://doi.org/10.1038/nbt.3942

Meng, W., Zhang, B., Schwartz, G. W., Rosenfeld, A. M., Ren, D., Thome, J. J. C., Carpenter, D. J., Matsuoka, N., Lerner, H., Friedman, A. L., Granot, T., Farber, D. L., Shlomchik, M. J., Hershberg, U., & Luning Prak, E. T. (2017b). An atlas of B-cell clonal distribution in the human body. *Nature Biotechnology*, *35*(9), 879–886. https://doi.org/10.1038/nbt.3942

Mesin, L., Ersching, J., & Victora, G. D. (2016). Germinal Center B Cell Dynamics. In *Immunity* (Vol. 45, Issue 3, pp. 471–482). Cell Press. https://doi.org/10.1016/j.immuni.2016.09.001

Mestas, J., & Hughes, C. C. W. (2004). Of Mice and Not Men: Differences between Mouse and Human Immunology. *The Journal of Immunology*, *172*(5), 2731–2738. https://doi.org/10.4049/jimmunol.172.5.2731

Metzger, T. C., & Anderson, M. S. (2011). Control of central and peripheral tolerance by Aire. In *Immunological Reviews* (Vol. 241).

Meyer, S., Woodward, M., Hertel, C., Vlaicu, P., Haque, Y., Kärner, J., Macagno, A., Onuoha, S. C., Fishman, D., Peterson, H., Metsküla, K., Uibo, R., Jäntti, K., Hokynar, K., Wolff, A. S. B., Meloni, A., Kluger, N., Husebye, E. S., Podkrajsek, K. T., … Hayday, A. (2016). AIRE-Deficient Patients Harbor Unique High-Affinity Disease-Ameliorating Autoantibodies. *Cell*, *166*(3), 582–595. https://doi.org/10.1016/j.cell.2016.06.024

Mora, T., & Walczak, A. (2016). *Quantifying lymphocyte receptor diversity*. http://arxiv.org/abs/1604.00487

Mostafavi, S., Yoshida, H., Moodley, D., Leboité, H., Rothamel, K., Raj, T., Ye, C. J., Chevrier, N., Zhang, S. Y., Feng, T., Lee, M., Casanova, J. L., Clark, J. D., Hegen, M., Telliez, J. B., Hacohen, N., de Jager, P. L., Regev, A., Mathis, D., & Benoist, C. (2016). Parsing the Interferon Transcriptional Network and Its Disease Associations. *Cell*, *164*(3), 564–578. https://doi.org/10.1016/j.cell.2015.12.032

Müller, L., Aigner, P., & Stoiber, D. (2017). Type I interferons and natural killer cell regulation in cancer. In *Frontiers in Immunology* (Vol. 8, Issue MAR). Frontiers Media S.A. https://doi.org/10.3389/fimmu.2017.00304

Nadkarni, S., Mauri, C., & Ehrenstein, M. R. (2007). Anti-TNF-α therapy induces a distinct regulatory T cell population in patients with rheumatoid arthritis via TGF-β. *Journal of Experimental Medicine*, *204*(1), 33–39. https://doi.org/10.1084/jem.20061531

Nagamine, K., Scott, H. S., Kudoh, J., Minoshima, S., Heino3, M., Krohn\, K. J. E., Lalioti4, M. D., Mullis, E., Antonarakis4, S. E., Kawasaki, K., Asakawa, S., To, F. I., & Shimizu, N. (1997). *Positional cloning of the APECED gene*. http://www.nature.com/naturegenetics

Nezos, A., Gravani, F., Tassidou, A., Kapsogeorgou, E. K., Voulgarelis, M., Koutsilieris, M., Crow, M. K., & Mavragani, C. P. (2015). Type I and II interferon signatures in Sjogren's syndrome pathogenesis: Contributions in distinct clinical phenotypes and Sjogren's related lymphomagenesis. *Journal of Autoimmunity*, *63*, 47–58. https://doi.org/10.1016/j.jaut.2015.07.002

Nurieva, R. I., Chung, Y., Martinez, G. J., Yang, X. O., Tanaka, S., Matskevitch, T. D., Wang, Y. H., & Dong, C. (2009). Bcl6 mediates the development of T follicular helper cells. *Science*, *325*(5943), 1001–1005. https://doi.org/10.1126/science.1176676

Ossart, J., Moreau, A., Autrusseau, E., Ménoret, S., Martin, J. C., Besnard, M., Ouisse, L.-H., Tesson, L., Flippe, L., Kisand, K., Peterson, P., Hubert, F.-X., Anegon, I., Josien, R., & Guillonneau, C. (2018). Breakdown of Immune Tolerance in AIRE-Deficient Rats Induces a Severe Autoimmune Polyendocrinopathy–Candidiasis–Ectodermal Dystrophy–like Autoimmune Disease. *The Journal of Immunology*, *201*(3), 874–887. https://doi.org/10.4049/jimmunol.1701318

Pereira, J. P., Kelly, L. M., Xu, Y., & Cyster, J. G. (2009). EBI2 mediates B cell segregation between the outer and centre follicle. *Nature*, *460*(7259), 1122–1126. https://doi.org/10.1038/nature08226

Peterson, P., Pitkänen, J., Sillanpää, N., & Krohn, K. (2004). Autoimmune polyendocrinopathy candidiasis ectodermal dystrophy (APECED): A model disease to study molecular aspects of endocrine autoimmunity. In *Clinical and Experimental Immunology* (Vol. 135, Issue 3, pp. 348–357). https://doi.org/10.1111/j.1365-2249.2004.02384.x

Piccoli, L., Campo, I., Fregni, C. S., Rodriguez, B. M. F., Minola, A., Sallusto, F., Luisetti, M., Corti, D., & Lanzavecchia, A. (2015). Neutralization and clearance of GM-CSF by autoantibodies in pulmonary alveolar proteinosis. *Nature Communications*, *6*. https://doi.org/10.1038/ncomms8375

Raybould, M. I. J., Rees, A. R., & Deane, C. M. (2021). Current strategies for detecting functional convergence across B-cell receptor repertoires. In *mAbs* (Vol. 13, Issue 1). Taylor and Francis Ltd. https://doi.org/10.1080/19420862.2021.1996732

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Ronkainen, M. S., Härkönen, T., Perheentupa, J., & Knip, M. (2005). Characterization of the humoral immune response to glutamic acid decarboxylase in patients with autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED) and/or type 1 diabetes. *European Journal of Endocrinology*, *153*(6), 901–906. https://doi.org/10.1530/eje.1.02026

Rothman, P., Chen, Y.-Y., Lutzker, S., Li,', S. C., Stewart, V., Coffman, R., & Alt', F. W. (1990). Structure and Expression of Germ Line Immunoglobulin Heavy-Chain £ Transcripts: Interleukin-4 Plus Lipopolysaccharide-Directed Switching to C. In *MOLECULAR AND CELLULAR BIOLOGY* (Vol. 10, Issue 4). https://journals.asm.org/journal/mcb

Rutz, S., & Ouyang, W. (2011). Regulation of interleukin-10 and interleukin-22 expression in T helper cells. In *Current Opinion in Immunology* (Vol. 23, Issue 5, pp. 605–612). https://doi.org/10.1016/j.coi.2011.07.018

Sansom, S. N., Shikama-Dorn, N., Zhanybekova, S., Nusspaumer, G., Macaulay, I. C., Deadman, M. E., Heger, A., Ponting, C. P., & Holländer, G. A. (2014). Population and single-cell genomics reveal the Aire dependency, relief from Polycomb silencing, and

distribution of self-antigen expression in thymic epithelia. *Genome Research*, *24*(12), 1918–1931. https://doi.org/10.1101/gr.171645.113

Saoudi, A., Kuhn, J., Huygen, K., Eur, al J., SaoudiO, A., KuhnO, J., HuygenO, K., de Kozak, Y., VeluA, T., GoldmanA, M., DruetO, P., & BellonO, B. (1993). TH2 activated cells prevent experimental autoimmune uveoretinitis, a TH1-dependent autoimmune disease. *Eur. J. Immunol*.

Segerberg, F., Lundtoft, C., Reid, S., Hjorton, K., Leonard, D., Nordmark, G., Carlsten, M., & Hagberg, N. (2019). Autoantibodies to Killer Cell Immunoglobulin-Like Receptors in Patients With Systemic Lupus Erythematosus Induce Natural Killer Cell Hyporesponsiveness. *Frontiers in Immunology*, *10*. https://doi.org/10.3389/fimmu.2019.02164

Setliff, I., Shiakolas, A. R., Pilewski, K. A., Murji, A. A., Mapengo, R. E., Janowska, K., Richardson, S., Oosthuysen, C., Raju, N., Ronsard, L., Kanekiyo, M., Qin, J. S., Kramer, K. J., Greenplate, A. R., McDonnell, W. J., Graham, B. S., Connors, M., Lingwood, D., Acharya, P., … Georgiev, I. S. (2019). High-Throughput Mapping of B Cell Receptor Sequences to Antigen Specificity. *Cell*, *179*(7), 1636-1646.e15. https://doi.org/10.1016/j.cell.2019.11.003

Simpson, E. (1949). Measurement of Diversity. *Nature*.

Sinclair, C., Bains, I., Yates, A. J., & Seddon, B. (2013). Asymmetric thymocyte death underlies the CD4:CD8 T-cell ratio in the adaptive immune system. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(31). https://doi.org/10.1073/pnas.1304859110

Stavnezer, J. (1996). Immunoglobulin class switching. *Current Biology*.

Stoeckle, C., Gouttefangeas, C., Hammer, M., Weber, E., Melms, A., & Tolosa, E. (2009). Cathepsin W expressed exclusively in CD8+ T cells and NK cells, is secreted during target cell killing but is not essential for cytotoxicity in human CTLs. *Experimental Hematology*, *37*(2), 266–275. https://doi.org/10.1016/j.exphem.2008.10.011

Street, K., Risso, D., Fletcher, R. B., Das, D., Ngai, J., Yosef, N., Purdom, E., & Dudoit, S. (2018). Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics*, *19*(1). https://doi.org/10.1186/s12864-018-4772-0

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P., & Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell*, *177*(7), 1888-1902.e21. https://doi.org/10.1016/j.cell.2019.05.031

Takemori, T., Kaji, T., Takahashi, Y., Shimoda, M., & Rajewsky, K. (2014). Generation of memory B cells inside and outside germinal centers. In *European Journal of Immunology* (Vol. 44, Issue 5, pp. 1258–1264). Wiley-VCH Verlag. https://doi.org/10.1002/eji.201343716

Thomas-Vaslin, V., Korthals Altes, H., de Boer, R. J., & Klatzmann, D. (2008). Comprehensive Assessment and Mathematical Modeling of T Cell Population Dynamics and Homeostasis 1. In *The Journal of Immunology* (Vol. 180). www.jimmunol.org

Turchaninova, M. A., Davydov, A., Britanova, O. v., Shugay, M., Bikos, V., Egorov, E. S., Kirgizova, V. I., Merzlyak, E. M., Staroverov, D. B., Bolotin, D. A., Mamedov, I. Z., Izraelson, M., Logacheva, M. D., Kladova, O., Plevova, K., Pospisilova, S., & Chudakov, D. M. (2016). High-quality full-length immunoglobulin profiling with unique molecular barcoding. *Nature Protocols*, *11*(9), 1599–1616. https://doi.org/10.1038/nprot.2016.093

Tuure Kinnunen. (2013). Accumulation of peripheral autoreactive B cells in the absence of functional human regulatory T cells. *Immunobiology*.

van den Berge, K., Roux de Bézieux, H., Street, K., Saelens, W., Cannoodt, R., Saeys, Y., Dudoit, S., & Clement, L. (2020). Trajectory-based differential expression analysis for single-cell sequencing data. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-14766-3

Vivier, E., & Malissen, B. (2005). Innate and adaptive immunity: Specificities and signaling hierarchies revisited. In *Nature Immunology* (Vol. 6, Issue 1, pp. 17–21). https://doi.org/10.1038/ni1153

Watson, C. T., & Breden, F. (2012). The immunoglobulin heavy chain locus: Genetic variation, missing data, and implications for human disease. In *Genes and Immunity* (Vol. 13, Issue 5, pp. 363–373). https://doi.org/10.1038/gene.2012.12

Wolock, S. L., Lopez, R., & Klein, A. M. (2019). Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Systems*, *8*(4), 281-291.e9. https://doi.org/10.1016/j.cels.2018.11.005

Yang, F., Nielsen, S. C. A., Hoh, R. A., Röltgen, K., Fabian Wirz, O., Haraguchi, E., Jean, G. H., Lee, J.-Y., Pham, T. D., Jackson, K. J. L., Roskin, K. M., Liu, Y., Nguyen, K., Ohgami, R. S., Osborne, E. M., Nadeau, K. C., Niemann, C. U., Parsonnet, J., & Boyd, S. D. (n.d.). *Shared B cell memory to coronaviruses and other pathogens varies in human age groups and tissues*. https://www.science.org

Yu, G., Li, F., Qin, Y., Bo, X., Wu, Y., & Wang, S. (2010). GOSemSim: An R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics*, *26*(7), 976–978. https://doi.org/10.1093/bioinformatics/btq064

Appendix

I. Supplementary material

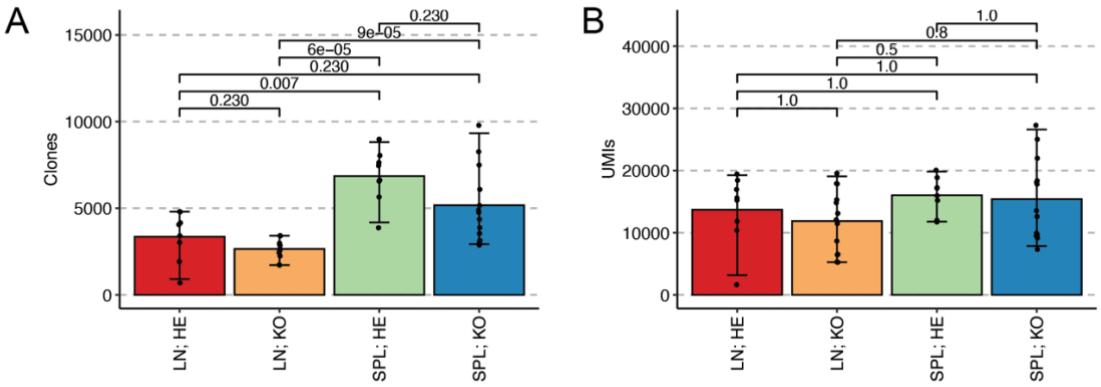

Figure S1. The number of clones and unique molecular identifiers (UMIs) detected in the Ig repertoire sequencing dataset. (**A**) The number of B cell clones (i.e. unique CDR3 sequences) in the dataset, grouped by tissue of origin and rat genotype. (**B**). The number of unique molecular identifiers (UMIs) that mapped to immunoglobulin sequences in the dataset. In both figures, error bars represent 25% and 95% confidence intervals. Wilcoxon rank-sum test was performed and Holm-Bonferroni-adjusted p values are indicated.

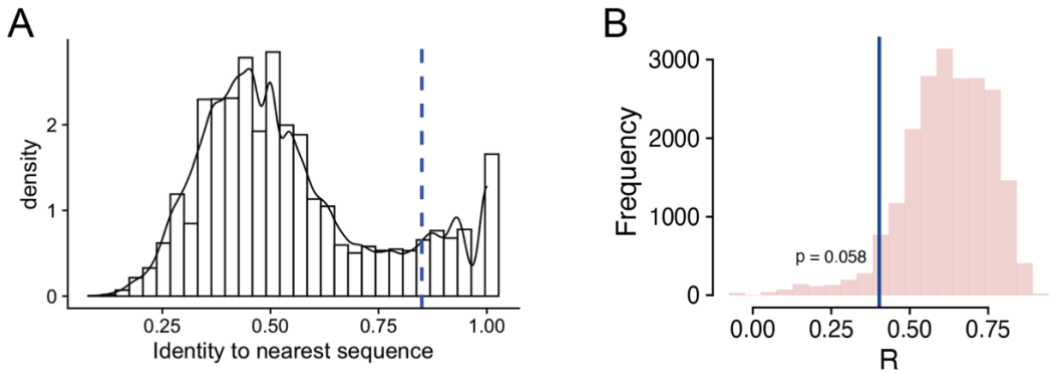

Figure S2. Sequence similarity cutoff for B cell lineage reconstruction and permutation test for correlation between IgG2b isotype frequency and age. (**A**) The distribution of pairwise sequence identity within groups of CDR3 sequences. The optimal cutoff for grouping B cells into clonotypes was determined by searching for the value of sequence identify that discriminated the two modes of the distribution. Sequence identity refers to normalized Levenshtein distance. (**B**) Permutation test to determine the difference between

correlation coefficients presented in **Figure 1C**. Because sample sizes are different in the two groups, we excluded 5 random *Aire-/-* lymph node samples and recomputed correlation coefficient (R). This procedure was repeated 10,000 times and the p-value was determined as the fraction of permutations that had correlation coefficient lower than that in the *Aire+/-* group (blue line).

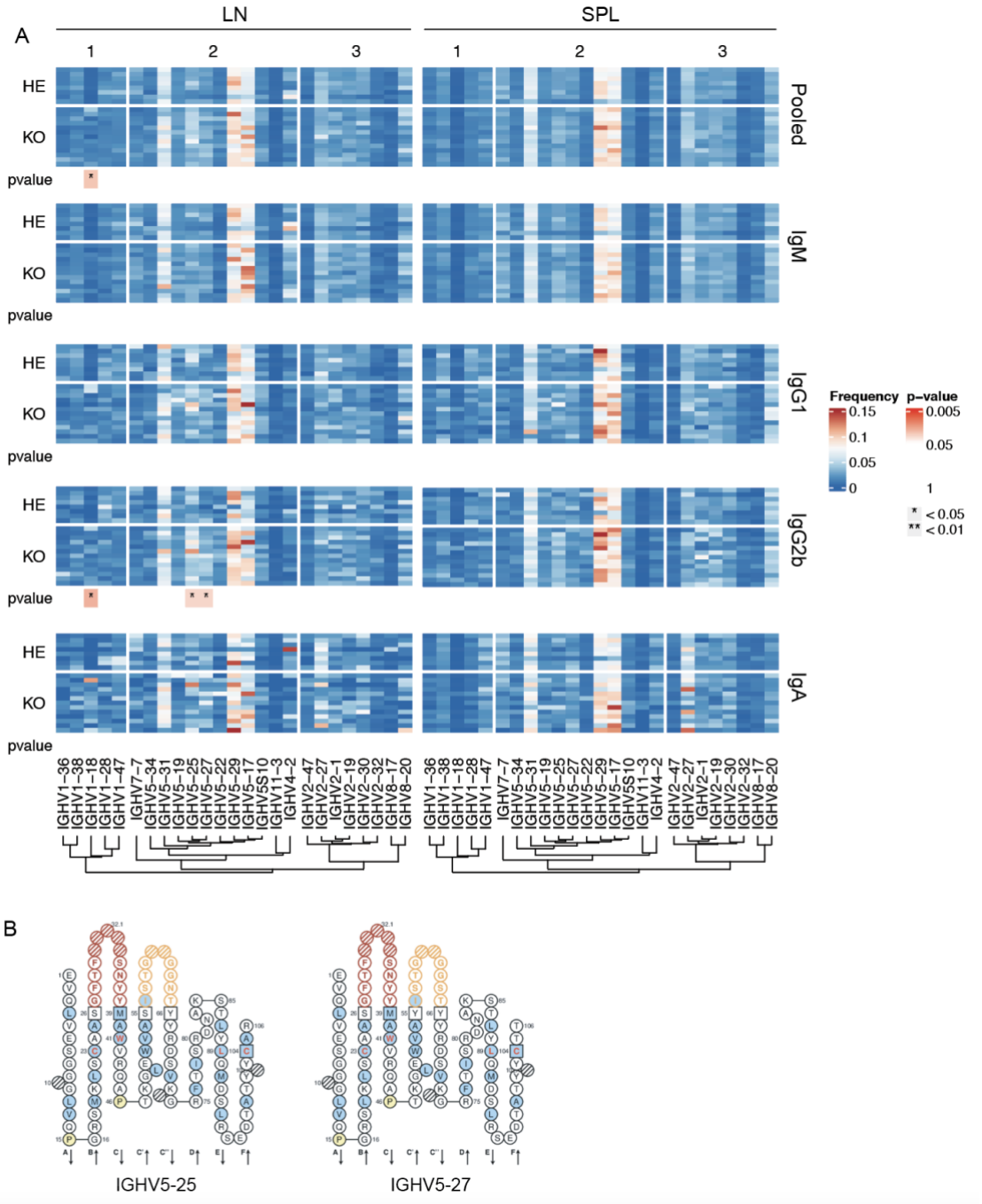

Figure S3. IGHV gene usage. (**A**) The frequency of IGHV genes in all major Ig subtypes (IgM, IgG1, IgG2b, IgA) separately and together (Pooled). Rows correspond to individual

samples and columns correspond to IGHV genes. Dendrogram at the bottom was obtained by performing average linkage clustering of V gene amino acid sequences and illustrates sequence similarity between different IGHV genes. ** $p \leq 0.01$, * $p \leq 0.05$, two-tailed Student's t-test with Holm Bonferroni correction. (**B**) A two-dimensional graphical representation of the variable domains encoded by IGHV5-25 and IGHV5-27 genes. Red circles correspond to CDR1 region, and orange circles correspond to CDR2 region, Adapted from IMGT (Lefranc et al., 2005).

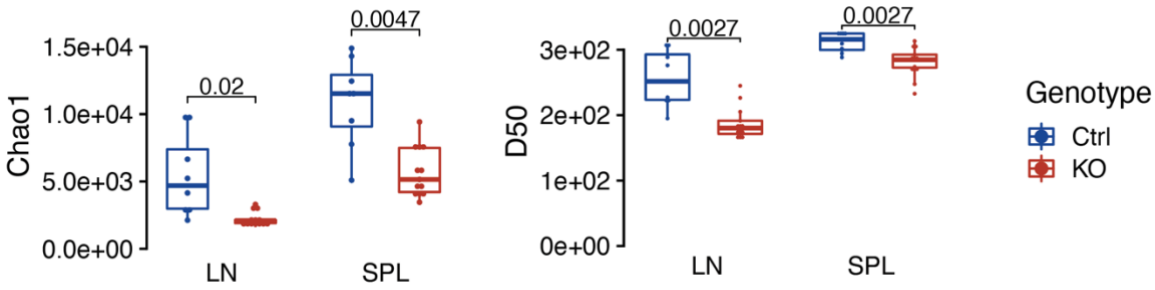

Figure S4. Immunoglobulin repertoire diversities in AIRE-deficient rats. Chao1 (**A**) and D50 (**B**) indices calculated on clonotype distributions. Holm-Bonferroni corrected p values from Student's t-test are shown.

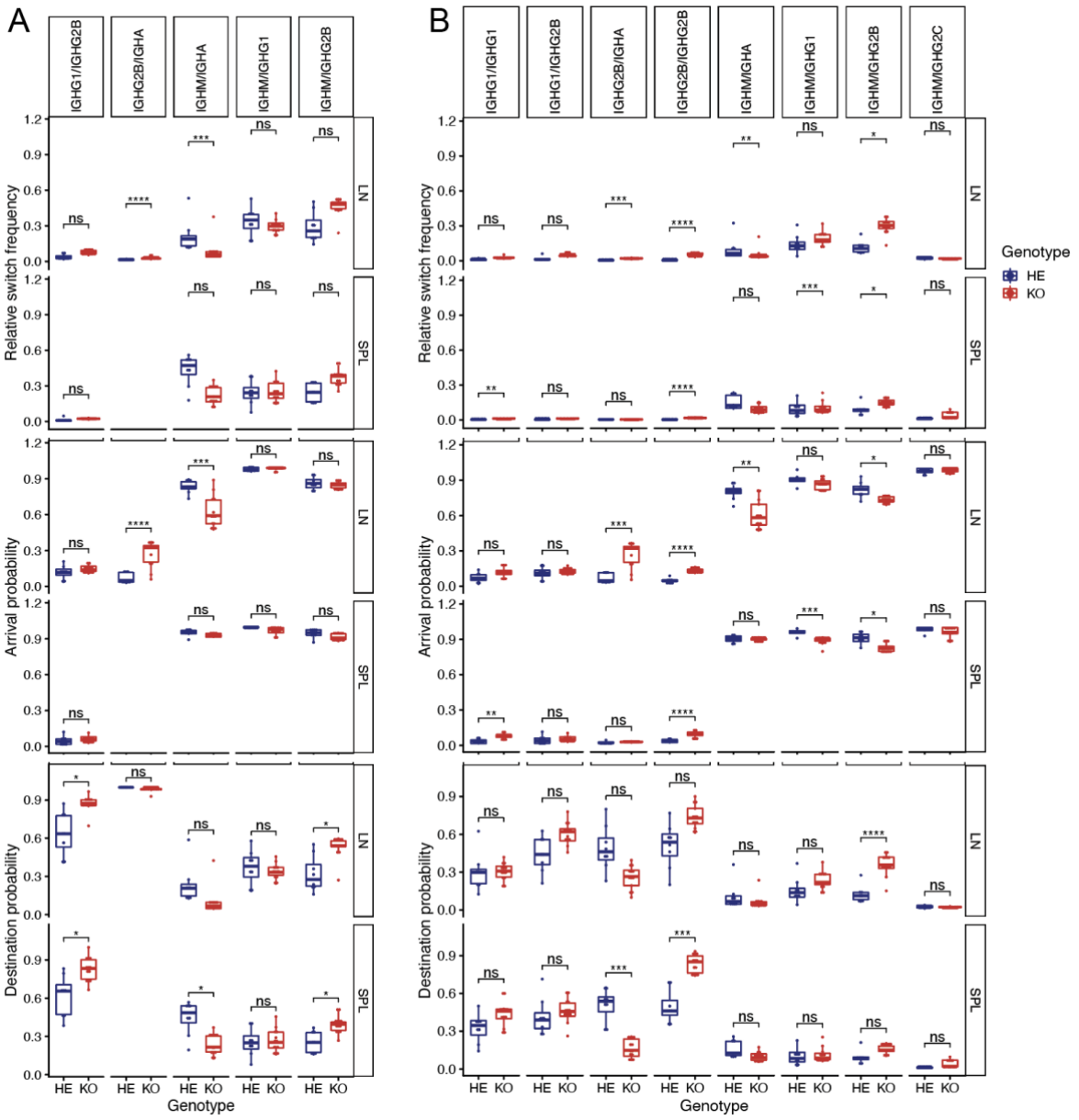

Figure S5. Class-switching landscape. Class-switch (**A**) probabilities and transition (**B**) probabilities were calculated as described in **Methods** (section 3.1.12). **** $P \leq 0.0001$, *** $p \leq 0.001$, ** $p \leq 0.01$, * $p \leq 0.05$, ns – not significant; Student's t-test with Holm Bonferroni correction.

Non-exclusive licence to reproduce the thesis and make the thesis public

I, Maksym Zarodniuk,
(*author's name*)

1. grant the University of Tartu a free permit (non-exclusive licence) to

reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright, my thesis

Profiling antibody responses and their contribution to disease in APS1 rat model using high-throughput sequencing approaches,
(*title of thesis*)

supervised by Ahto Salumets and Prof. Pärt Peterson.
(*supervisor's name*)

2. I grant the University of Tartu a permit to make the thesis specified in point 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 4.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in points 1 and 2.

4. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Maksym Zarodniuk
27/05/2022