ABSTRACT

Identification and Analysis of an Improved Animal Model for Behçet's Disease: A Computational Approach

Samuel J. Shenoi

Director: Dr. Erich Baker, Ph.D

Behçet's disease (BD) is a multi-system inflammatory disease primarily affecting patients along the historical silk road. Current attempts to develop an animal model for BD have proved to be only partially successful; they have been unable to replicate both the phenotypic and genotypic aspects of this complex disease. In order to overcome these limitations, we performed a computational analysis to find an improved animal model for BD. First, we identified phenotypic and genetic factors of BD. Second, we compared these factors to 27 other autoimmune diseases. Finally, we analyzed identified autoimmune diseases animal models in order to determine if they can be used for BD. Using this methodology, we were able to identify 3 potential animal models for the continuation of BD research. These three models may prove useful in the development of novel BD treatments.

APPROVED BY DIRECTOR OF HONORS THESIS:

Erich J Baker

Dr. Erich Baker, Department of Computer Science

APPROVED BY THE HONORS PROGRAM:

Dr. Andrew Wisely, Interim Director

DATE: _____

IDENTIFICATION AND ANALYSIS OF AN IMPROVED ANIMAL MODEL FOR BEHÇET'S DISEASE: A COMPUTATIONAL APPROACH

A Thesis Submitted to the Faculty of

Baylor University

In Partial Fulfillment of the Requirements for the

Honors Program

By

Samuel J. Shenoi

Waco, Texas

May 2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Acknowledgements

I would like to extend a special thanks to the following people for their help with this work. Dr. Timothy Reynolds provided the code base used for the VZ subway plot visualization tool used in Chapter 2. Frimpong Baudo assisted with the setup of the Graph Autoencoders for Chapter 2. Dr. Jason Bubier provided access to The Jackson Laboratory's Human Mouse Disease Connection website used for Chapter 4. Dr. Mary Lauren Benton provided guidance for the modified Jaccard formula used in Chapter 3. Finally, Candice Ditsch proofread select chapters of this thesis. Thank you once again for all of the help that you provided during the duration of this work.

I would also like to sincerely thank Dr. Mary Lauren Benton and Dr. Elissa Chesler for serving as members of my thesis defense committee and for taking the time to read this thesis in its entirety.

Finally, I would like to especially thank my thesis director, Dr. Erich Baker, for his support and guidance throughout the entirely of this project.

INTRODUCTION

Globally, BD affects males and females at roughly equal rates, and the onset of the disease occurs in the third decade of life [1]. Currently, the disease is classified as an autoimmune or auto inflammatory disease and there exists no cure, though recent research has resulted in the push to reclassify BD a a syndrome rather than a disease [1, 2]. The disease is thought to be triggered in genetically predisposed individuals by a combination of viral infection and environmental factors, but the exact mechanism of infection is not currently known [3].

Developing accurate animals models representing the disease is vital for the discovery of the pathogenesis and potential treatments for the disease. Future animal models need to be able to reflect both clinical presentation and underlying genetic features of the disease. Current BD animal models have proved to be only partially successful as they are often unable to replicate both the phenotypic and genotypic aspects of the disease [4, 5, 6].

In this work, we undertake a computational approach towards finding an improved animal model for BD that is able to account for both genetic and environmental factors of the disease. To do this, we leverage the computational power of the Geneweaver HiSIM graph tool. The HiSIM Graph tool, previously called The Phenome Map tool, on Geneweaver results in a graph that is a hierarchical network of multiway gene set intersections. The resulting geneset interaction graph enables users to find genes connected to all populated subsets of an input set of gene lists[7]. To accomplish this task, the tool takes takes advantage of bipartite data structure in order to dynamically enumerate maximal bicliques to be arranged into hierarchical associations [8, 7].

We complete our computational approach in 3 stages: 1) the identification of genetic and phenotypic factors of BD, 2) the comparison of the genetic factors of BD to the genetic factors of other autoimmune disease, and 3) the comparison animal models from identified autoimmune disease to the phenotypic factors of BD. In order to augment our

analysis, we introduce improvements to the HiSIM graph visualization that will more effectively display geneset interactions.

CHAPTER 1

Behçet's Disease Review

1.1 Introduction

As previously mentioned, developing accurate animal models representing BD is vital for the discovery of pathogenesis and potential treatments for the disease. Current BD animal models have been unable to replicate both the phenotypic and genotypic aspects of the disease [4, 5, 6], which yields inconsistent and incomplete disease manifestation. Future animal models may overcome these inefficiencies only if they are able to reflect both phenotypic presentation and underlying genetic features of the disease. Therefore, in our quest to identify an improved animal model for the disease, we outline the phenotypic presentation, genetic influences, and their impact on the current animal model landscape of BD in this chapter.

1.2 Phenotypic Presentation

1.2.1 Symptoms

BD was first described as a multi-system disease demonstrating lesions of the mouth, the genitalia, and the eye [9]. Symptoms of the disease manifest at different points in the patients life with some symptoms, such as oral aphthosis, presenting earlier than other symptoms [10].

Oral aphthosis, a well-defined painful round or oval ulceration with a white base surrounded by a red areola, is the most common symptom of the disease [3, 1]. These lesions are the first typical phenotypic indications of the disease and occur in 86 - 100% of patients [3]. Oral lesions vary in size from a few millimeters to centimeters and are localized at the lips, buccal mucosa, soft palate, and tongue of patients [11].

Genital aphthosis is the second most common symptom of the disease with 57% - 93% of BD patients reporting this symptom. These ulcers are larger but otherwise similar to oral ulcers present in BD patients. They are found most commonly on the scrotum of men and the vulva in women but can also be found on the epididymis and penis of men and vagina and cervix in women [3].

Skin conditions such as pseudofolliculitis (also called Behçet's pustulosis), and erythema nodosum are commonly found in 60% - 80% of patients. Skin aphthosis occurs more rarely in patients [1]. Pseudofolliculitis lesions, resembling acne, are found over the range of the body and are not always hair follicle based. Erythema nodosum lesions occurs as bilateral, pretibial, painful, and hot erythematous nodules. They are commonly found localized to the face, neck, buttocks, and forearms [12].

Ocular lesions are another common symptom of the disease with 40% to 60% of patients presenting with these symptoms [1, 3]. Manifestations of this symptom are usually present in both eyes and occur 2-3 years after symptoms start. Posterior uveitis occurs in almost all cases of ocular involvement while anterior uveitis occurs less frequently [12].

Other less common symptoms of the disease include nervous system manifestation and gastrointestinal involvement. These symptoms along with vascular manifestations are considered special type Behcet's disease[10].

Neuro-Behçets Disease(NBD) usually is present in the Central Nervous System and is further divided into parenchymal NBD and non-parenchymal NBD. Parenchymal NBD presents with one or more of the following: ophthalmoparesis, cranial neuropathy, cerebellar or pyramidal dysfunction, cerebral or spinal cord involvement, myelopathy, encephalopathy, hemiparesis, hemisensory loss, seizures,dysphasia, mental changes including cognitive dysfunction, and psychosis. Non-parenchymal NBD presents with cerebral venous thrombosis, intracranial hypertension syndrome, or acute meningeal syndrome [13].

Gastric manifestations of BD commonly include esophageal ulcers, gastric ulcers and/or duodenal ulcer, small and/or large intestinal ulcers, large artery aneurysm/thrombosis in abdomen, and large vein thrombosis in the abdomen [14].

1.2.2 Physiological Manifestations

Natural Killer(NK) cells play a cytotoxic role in infected cells and tumor cells but also regulate the function of other immune cells, including dendritic cells (DCs) and T cells, by secreting cytokines. It has been reported that patients with BD have less NK cells in their peripheral blood, and NK cells are implicated to play an important role in Th1 dominance in BD patients [15].

Patients presenting with BD have been shown to have an increase of neutrophils usually involved in perivascular infiltration in BD lesions [15]. Neutrophils, one of the first lines of defense against invading microbes, can damage host cells and tissues while destroying microbes. Tissue damage is one of the main triggers of inflammation, which in turn triggers the immune response. Neutrophil hyperactivity is stimulated by antigen-presenting cells (APCs) and results in the stimulation of T helper 1 (Th1) response with neutrophil secreted cytokines [3].

1.2.3 Diagnostic Criteria

Current diagnostic criteria of BD follows the International Criteria for Behçet's Disease (ICBD) guidelines [16]. The diagnostic criteria for BD has undergone 6 iterations with ICBD being the 6th iteration at the time of writing [17]. ICBD diagnostic criteria relies on a point-based system with ocular lesions, oral aphthosis, and genital aphthosis each being assigned 2 points. Skin lesions, central nervous system involvement, vascular manifestations, and the, when used, pathergy test are all assigned 1 point. Patients scoring 4 or more points using this criteria are classified as having BD [16].

Separate diagnostic criteria exist for BD subtypes. Neuro-Behçets disease diagnosis requires a BD diagnosis, a neurological syndrome recognized to be associated with BD, and no better explanation of neurological findings[13]. Diagnosis of gastrointestinal BD is based mainly on clinical findings and should be considered only after ruling out Crohn's disease (CD), tuberculosis, vasculitis, and other gastrointestinal diseases [14].

1.3.1 HLA-B51 and other HLA genes

The Human Leukocyte Antigen B51 (HLA-B51) has long been thought to be associated with Behcet's disease [12, 18]. HLA-B51 is a variation of the HLA-B gene, which is part of a family of genes called the human leukocyte antigen (HLA) complex. These genes are components of the human version of the major histocompatibility complex (MHC) and the HLA-B gene is further classified as a MHC class I gene [19]. The function of molecules coded for by MHC genes is to bind peptide fragments derived from pathogens and display them on the cell surface for recognition by the appropriate T cells [20]. These T cells then work to rid the body of the pathogen.

In virtually every studied population, HLA-B51 has been the most closely associated risk factor for BD [21]. Studies analyzing the genetic influences of BD in distinct populations - Iranian, Japanese, Italian, Native American, Hispanic, Israeli, and Irish - have corroborated HLA-B51's association with BD [22, 23, 24, 25, 26, 27, 28]. Further studies have shown the prevalence of HLA-B51 in BD patients to range between 50% - 72% compared to only 10-15% in healthy controls [29, 30]. The exact mechanism behind the association of HLA-B51 and BD is still unknown. Possible theories of HLA-B51's causative effect on Behcet's disease include the presentation of cytotoxic CD8+ T cells and the interaction of HLA-B51 with NK cells [31].

Verity *et al.*'s report on the geographic distribution of HLA-B51 shows a correlation between the prevalence of HLA-B51 in a population and the prevalence of BD. Populations presenting with a high prevalence of HLA-B51 lie along the historic silk road mimicking the distribution of BD, the "Silk Road Disease". However, a sizable portion of BD patients present as HLA-B51 negative and the high prevalence of HLA-B51 in American Indian populations where BD is rare indicates the role of other genetic loci as factors of BD pathogenesis [21, 26]. Research has also found a higher frequency of HLA-B51 in males compared to females and there does not appear to be an association between HLA-B51 and the severity of the disease. [28, 32]. It is for these reasons

that, despite a strongly confirmed association, HLA-B51 has yet to be classified as the primary causative agent of BD.

There have been reports of other HLA genes being associated with BD. As mentioned earlier, HLA genes are components of the human version of the MHC, and the function of molecules coded for by MHC genes is to bind peptide fragments derived from pathogens and display them on the cell surface for recognition by the appropriate T cells [20, 19]. HLA-A02, HLA-A24, HLA-A26, HLA-A31, HLA-B27, and HLA-B57 were found to increase susceptibility to BD [33, 34]. HLA-A03, HLA-B15, HLA-B35, HLA-B49, HLA-B58 were reported as BD-protective [33]. HLA-A29 has been shown to be increased in patients with ocular symptoms when coupled with a decrease in HLA-B51. HLA-B35 frequency was shown to be decreased in patients with thrombophlebitis. Finally, HLA-B3901, Cw1, Cw14, Cw15, Cw16, HLA- DRB104, and HLA-DRB107 have also been described to be associated with BD[34].

1.3.2 TNF and TNFAIP3

Previous studies have indicated an association between Tumor Necrosis Factor(TNF)-α and BD[12, 35, 36, 37]. TNF-α is a pro-inflammatory cytokine that is important in the regulation of the immune response [12, 34, 38]. TNF is encoded in the class III region of the HLA complex adjacent to HLA–B which provides some positional evidence for its involvement in BD [37]. Specifically, research has indicated a relationship between the TNF-α Induced Protein 3 (TNFAIP3) gene and BD [39, 33, 40, 41].

Research into the particular variants of TNF-α that confer BD susceptibility have had mixed results. Radouane *et al.* conducted a study in a Moroccan population that found an association with the -1211C allele and BD [42]. Lee *et al.* found no difference in the distribution of TNF-α -308G/A polymorphisms in Korean BD patients versus healthy controls [43]. Touma *et al.*'s 2010 meta-analysis found significant associations between -1031C, -857T, and the -238A alleles with BD while also finding no association between -863A and -308A/G alleles [44]. However, in 2013 Zhang *et al.* conducted a meta-analysis of BD polymorphisms that confirmed the association between TNF308A/G and

BD susceptibility while also identifying significant associations between TNF 238A/G, 1031C/T, and 857T/C polymorphisms and BD risk[45]. It is speculated that the apparent contradictory results partially stem from variation among ethnic groups tested[44].

1.3.3 MICA

MHC class I chain-related gene A (MICA) has a speculative relationship to BD [12, 36]. It is located between between the HLA-B and TNF genes[18] and is associated primarily though genomic proximity.

Research into MICA's effect on BD has had mixed results. In a Japanese population, an association between A6, a MICA-transmembrane region, and BD was discovered [46, 47]. These findings were later corroborated in Greek, Korean, Turkish, and pediatric Italian populations [48, 49, 50, 51, 47]. However, follow-up studies seem to indicate that these findings were due to the linkage disequilibrium (LD) effect of HLA-B51 rather than being caused by MICA-A6 [47]. Additional evidence for the LD effect theory came from a study in an Italian population that identified a strong association between HLA-B51 and MICA, but was unable to identify a relationship between MICA and BD [25].

These conflicting results were then tested in meta-analyzes of previous case studies which found that a correlation existed between BD and MICA-A6 in both Caucasian and Asian populations [52, 53]. For this reason, MICA is thought to be a gene of interest in BD.

1.3.4 ERAP1

Endoplasmic reticulum-associated amino-peptidase 1 (ERAP1) has been identified as potential gene of interest in BD susceptibility in patients presenting with HLA-B51 [54, 10, 55, 56, 57]. The ERAP1 gene is located on chromosome 5q15 and encodes an amino-peptidase responsible for N-terminal trimming of peptides in the endoplasmic reticulum. This is a critical step of the processing of the peptides to optimize their length for MHC-I binding [56].

ERAP1 has been associated with ankylosing spondylitis (AS) and psoriasis, two

other immune-mediated diseases associated with MHC class I genes [10, 56, 58]. Furthermore, an association exists between these diseases and the IL-23R and IL10 genes, two other BD associated genes, which might indicate the existence of shared pathogenic pathways [56].

ERAP1 is a highly polymorphic gene, and studies into the particular variants of ERAP1 that confer BD susceptibility have yielded some results. The most common variants, designated as haplotypes (Hap), or combinations of non-synonymous Single Nucleotide Polymorphisms (SNPs), are named Hap1 to Hap10. Together these 10 haplotypes account for 99% of the natural ERAP1 variants in human populations[59, 60]. Specifically, Hap10 has been shown to induce extensive changes in the peptide repertoires available for loading onto HLA-B*51, which alter its antigen-presenting specificity and can destabilize the molecule through generating a peptidome of lower affinity [59]. This haplotype has been shown to induce the most changes that negatively affect ERAP1 activity [60].

1.3.5 STAT4

STAT4, a signal transducer and activator of transcription, is another gene that has been correlated with BD [61, 33, 12]. STAT4 activates gene expression involved in functional regulation and differentiation of T-helper cells, NK cells, mast cells and dendritic cells, and it modulates differentiation of naïve T cells into Th1 and Th17 cells[33].

Genome Wide Association Studies (GWAS) conducted in individuals of Han Chinese descent showed a susceptibility locus around STAT4 SNPS rs7574070, rs7572482, and rs897200 providing evidence for STAT4's involvement with BD [62]. The results linking STAT4 SNP rs7574070 to BD were corroborated in a study in an Iranian population [57]. Kim *et al.* conducted a study on Intestinal BD in a Korean population that looked at STAT4, IL17A, and IL23R. Their study was unable to find a significant difference in the genotype frequencies of SNPs in STAT4 between the patients and controls, but found associations between G149R in IL23R and rs11685878 in STAT4, rs2275913 in IL17A and rs7574865 in STAT4, and rs1188941 in STAT4 and rs2275913 in IL17A

[63].

1.3.6 IL

Multiple members of the interleukin family of genes have been implicated in BD [33]. Interleukins (IL) are cytokines that play essential roles in the activation and differentiation of immune cells. They have pro-inflammatory and anti-inflammatory properties, and their primary function is to modulate growth, differentiation, and activation during inflammatory and immune responses[64].

An association between IL-10 and BD has been identified in multiple independent studies [65, 66, 67, 68, 69]. IL-10, a member of the interleukin family, has a variety of functions including decreasing the antigen presentation and MHC class II expression of dendritic cells, decreasing co-stimulatory molecule B7-1/B7-2 expression on monocytes and macrophages, and limiting the production of proinflammatory cytokines[64, 70]. The pathological role and functional basis of the gene is yet to be determined. Research into IL-10 found that the gene's effect on BD is primarily localized to the impaired migration of M2 macrophages into inflammatory lesions, but the exact mechanism of action is still unknown [71]. A meta-analysis studying IL-10 demonstrated significant BD associations with IL-10 polymorphisms -819C/T, 592C/A, rs1518111G/A, and rs1554286C/T and found that these polymorphisms may play protective roles against BD. These studies also found that the -1082A/G polymorphism carried no significant association with BD [72, 73].

Likewise, an association between IL-12 an BD has been previously described [33, 74, 69, 68]. IL-12 is produced by monocytes, and it causes induction of Th1 cells while also being a potent inducer of interferon gamma production by T lymphocytes and NK cells [64, 75]. Aridogan *et al.* found patients presenting with Recurrent Aphthous Stomatitis (RAS), a common oral mucosal disease with a similar immunopathogenesis to BD, had higher levels of IL-12 compared to controls. A decrease in IL-12 was found in active BD patients compared to inactive BD patients. These findings were corroborated in a study that looked at the sera of inactive, mildly active, and active BD patients [68, 69].

Studies have shown that Heat Shock Proteins (HSPs) stimulate T cells and thus affect the release of proinflammatory cytokines and IL-12[75, 74]. Finally, advances in IL-12-family related genes have revealed two major subclusters - Th17/Th1 and Th1/IL-35 - which have been associated with multiple other diseases including inflammatory bowel diseases, psoriasis, ankylosing spondylitis, rheumatoid arthritis, primary biliary cirrhosis, and Graves' disease. Some research has shown that BD might be included in the Th17/Th1 cluster [76].

Multiple studies have indicated a relationship between IL-23 and BD [33, 77, 78, 79]. IL-23 impacts the development of Th17 cells [33]. IL-23 also promotes IL-17 production which, along with TNF, induces a sustained neutrophil recruitment during inflammation [80, 78].

Other IL family members, including IL-6 and IL-11, have also been shown to have generalized correlation with BD [81, 82, 33]. However, the roles of these particular cytokines have very limited exploration.

1.3.7 Inflammatory Genes

Genes associated with inflammation including CIITA,MEFV,IRF8,REL,TLR2,4, CCR1-CCR3, GIMAP, KLRC4, and interferon-γ have been shown to be associated with Behcet's disease [33, 83, 3]. Research looking into CCR1 found that it's effect on BD might be primarily localized to impaired migration of M2 macrophages into inflammatory lesions, similar to the IL-10 mechanism [71].

Other genes including NCOA5, FOXP3, PSORS1C1, FUT2, UBAC2, SUMO4, Loci at ADO-EGR2, CEBPB-PTPN1, JRKL-CNTN5 have been shown to confer BD susceptibility [33]. Finally, there have been conflicting reports of the NEMO, NOD2, PSTPIP1, MVK and TLR4 genes having associations with BD. Further research is needed to clarify the role, if any, these genes play in BD pathogenesis [12, 83].

1.4 Animal Models

To date there have been multiple attempts to create an animal model of BD, resulting in varying levels of success [4, 5].

1.4.1 Environmental Model

Initial models of BD used chemicals to induce symptoms of BD in Pitman-Moor swine[84, 4]. The chemicals were administered orally and included organic chlorides, organo-phosphate (DDT-trichloroethanediyl-bis-chlorobenzene, polychlorinated-biphenyl (PCB), Sumithion™ – dimethyl-nitro-phosphorothioate), and inorganic copper [84]. All animals eventually developed symptoms including folliculitis, cutaneous nodules, genital ulcers, oral aphthae, and intestinal ulcers[4, 84]. However, the time required to cultivate these symptoms, over a year, made this model infeasible for future studies [4].

1.4.2 Herpes Simplex Virus Model

The primary animal model of BD involves the inoculation of mice with Herpes Simplex Virus (HSV) [85]. BD has been long thought to have a viral triggering factor; Hulusi Behcet described viral artifacts present in patients during his first paper describing the disease [9]. Subsequently, HSV was found to be associated with BD [86, 87]. In creating HSV-induced BD, HSV type 1 (KOS strain) is inoculated into ICR mice through a scratched earlobe resulting in oral, genital and skin ulcers, eye lesions, and arthritis which mimics symptoms observed in BD patients [88, 85]. In the original paper describing the process, 33% of mice died after infection, 29.8% showed symptoms, and the rest retained a normal appearance. Of the mice presenting symptoms, 57% presented with skin ulcers, 39% presented with eye syndromes, and 19.5% presented with genital ulcers among other symptoms [85].

This model's limitations have been outlined in Charles *et al.* and Sohn *et al.* [6, 85]. Briefly, only 1/3 of the mice infected by this protocol displayed symptoms making very large sample sizes a requirement to effectively use this model [6, 85]. Additionally, the

lack of input of genetic and environmental factors in the development of the model might undercut it's effectiveness as an animal model of the disease since BD is thought to be triggered by a combination of a genetic and environmental factors [87]. A proposed modification to this model in order to account for underlying genetic factors is to infect transgenic HLA-B51 mice with HSV [6, 87, 89]. To our knowledge this model has yet to be tested.

1.4.3 Transgenic HLA-B51 Models

Attempts to create a transgenic mouse model presenting with the HLA-B51 allele have proved only partially successful. Takeno *et al.* created a transgenic mouse model presenting with the HLA-B51 allele. However, while the mice displayed an increase in the neutrophil production, consistent with BD in humans, they did not present any of the clinical symptoms of BD [89]. Karaki *et al.* demonstrated that HLA-B51 transgenic mice can be useful in the production of allospecific anitbodies for HLA-A and HLA-B antigens [90]. Finally, Sato *et al.* established HLA-B*51:01 transgenic humanized mice and demonstrated the elicitation of human effector CD8+ T cells in the transgenic mice, but not in controls, after infection with HIV-1 [91]. These findings might provide evidence for the effectiveness of infecting HLA-B51 transgenic mice with HSV since infection of a virus was able to elucidate this response [6, 87, 89, 91, 3]. The utility of this approach, however, remains speculative to date.

1.4.4 Heat Shock Protein Model

Heat shock proteins have been reported in patients presenting with BD and Intestinal BD, and may present yet another potential trigger for the disease[92, 93, 75]. The development of a mouse model presenting Uvelitis-like symptoms was achieved in Lewis rats using peptides 336-351 derived from human 60-kDA HSP. Uveitis was observed in 80% of rats immunized with a combination of this peptide and B. pertussis[6]. Unfortunately, no other symptoms of BD were identified in these mice [94, 92].

1.4.5 Streptococcal Model

Supporting the notion that Behçets disease may have a bacterial or viral triggering factor [9], members of the Streptococcal family have been observed in oral ulcers of patients presenting with BD in elevated levels. Furthermore, experiments in human dermal microvascular endothelial cells have shown that *S. sanguinis* can induce a similar inflammatory response as BD patient sera [95, 96, 97]. While this has led to the hypothesis that the Streptococcal family could play a role in BD pathogenesis [96, 98], attempts to develop animal models based on this information remain wholly unsuccessful. Animal experiments utilizing *S. salivarius*, *S. faecalis*, *S. pyogenes*, or *S. sanguisor* induced acute multiorgan infectious/inflammatory reactions, septic shock, and noninfiltrative short-term uveitis. It is therefore unsurprising that the inability of these models to produce an experimental model of BD has led to the lack of recognition of the streptococcal model and the search for other bacterial components for the development of a bacterial model [99, 5, 84].

1.4.6 α-tropomyosin model

It has been shown that the sera of BD patients contain IgG antibodies directed towards α-tropomyosin protein [84, 100, 101]. Two models have been created based on this information. The first model immunized Lewis rats with bovine α-tropomyosin in CFA injected into the hind foot. The rats immunized by this methodology developed anterior uveitis and inflammatory skin infiltrates [100, 6]. The other model inoculated Lewis rats with bovine alpha-TPM and differed from the first model only in the site of injection[84]. Ninety percent of the immunized rats developed severe acute arthritis 12 days after vaccination, and rats that were followed-up for 6 months had persistent inflammation of the leg joints. These rats did not display uveitis symptoms [102].

1.4.7 Human Sera Model

A recent model of Neuro-Behçet's disease (NBD) inoculates Sprague-Dawley rats with IgGs isolated from the sera of neuropil antibody-positive NBD patients [103, 6].

This model was able to induce a negative effect on locomotor activity and a neurotoxic action. However, no other appreciable clinical symptoms or death in rats were identified [103]. Support for this model comes from the fact that the infectious agent used was obtained from NBD patients, supporting it as a triggering factor of the disease [6]. Additionally, the fact that serum IgG, but not IgG depleted sera, was shown to induce the response lends evidence to the classification of BD as an autoinflammatory disease [6, 103].

1.4.8 Induced Uveitis models

The final group of models used for BD research are induced models of Uveitis. Several animal models of Experimental Uveitis have been described, some of which have been employed towards BD research. The most notable ones prescribed for BD research include interphotoreceptor retinoid-binding protein (IRBP) in BIO.A, BlO.S (8R), A, and MRL (+/+) mice, recoverin in Lewis rats, and retinal soluble antigen or S-Ag in SJL/J X AKR, and AKR mice[6, 84]. Of particular interest in the identification of these models was the fact that the S-Ag epitope shared homology to a conserved sequence in the HLA-B molecules [84]. While the downsides of these models lie in the fact that they only are able to replicate Uveitis symptoms, they present a valuable opportunity to research BD symptoms in isolation [6].

1.5 Conclusion

Behçets disease (BD) is a multi-system inflammatory disease presenting in patients along the historic Silk Road [1]. Developing accurate animal models representing BD is vital for the discovery of pathogenesis and potential treatments for the disease. The ideal animal model should replicate, to a major extent, both a human disease phenotype and its underlying causality [104]. To date, current animal models of BD have only been partially successful in achieving this ideal as they present exclusively with either the phenotype or the genotype of the disease [4, 5, 6].

The difficulty of developing an animal model for BD is due, in large part, to the inherent difficulties of developing animal models for complex, multi-system diseases. Despite the vast similarities that exist between humans and other mammals, significant differences exist across molecular, physiological and developmental scales [105]. Moreover, the recent push to reclassify of BD as a syndrome rather than a disease underscores additional challenges in producing a consistent animal model. Observations that distinct subsets of BD display distinct expression profiles and different disease associated pathways [2] raises the possibly that an animal model for one BD subset will be ungeneralizable to other BD subsets.

Despite these challenges, hope exists for discovery of an improved animal model for Behçet's Disease. Depression, arguably a much more difficult disease to replicate in animal models, research has uncovered animal models that are able to replicate both genetic and phenotypic features of the disease [106, 107]. While these models are far from perfect representations of the human version of the disease, their existence illustrates that it is possible to develop approximate models capable of replicating phenotype and underlying causality [104].

CHAPTER 2

HiSIM Graph Considerations

2.1 Introduction

After identifying the genetic factors of BD in the preceding chapter, we then sought to improve the Geneweaver HiSIM tool visualization that would be used to analyze these factors. The Geneweaver HiSIM Graph tool results in a bipartite graph that is a hierarchical network of multiway gene set intersections [8, 7]. The bipartite nature of the HiSIM tool presented a specific opportunity for the improvement of the HiSIM graph visualization. Since a bipartite graph is a set of graph vertices that can be split into two sets such that no two graph vertices within the same set are connected by an edge [108, 109], the graph permitted the hierarchical display of geneset interactions as HiSIM graph tool output and allowed for the adoption of subway plots to be used as a visualization technique.

However, integrating the subway plot visualization proved challenging due to the inability of certain graphs to be displayed by the subway plots. Therefore, we load tested the subway plot visualizations in order to determine the conditions that caused graphs to not display. In this chapter, we outline 1) the integration of the HiSIM plots into the Geneweaver ecosystem and 2) the load testing of these plots.

2.2 Methods

2.2.1 VZ Integration

In order to better visualize the Geneweaver HiSIM graph, a Subway plot visualization (VZ) tool was adapted to work within the exisiting Geneweaver ecosystem [1]. The VZ

[1] https://github.com/treynr/vz

subway plots were implemented in d3.js which is a representation transparent approach for the web. D3 allows designers to bind input data to document elements making it ideal for representing dynamically created graphs [110].

In order to integrate VZ into the existing Geneweaver ecosystem, a new python module was created. The python module transposed the input data generated by the Geneweaver HiSIM tool into a format that was compatible with the VZ subway plots. The existing view on Geneweaver was updated in order to allow users to switch between visualizations of the HiSIM graph. Finally, a minified version of the Geneweaver website focused solely on the HiSIM graph was created for testing purposes.

2.2.2 Load Testing

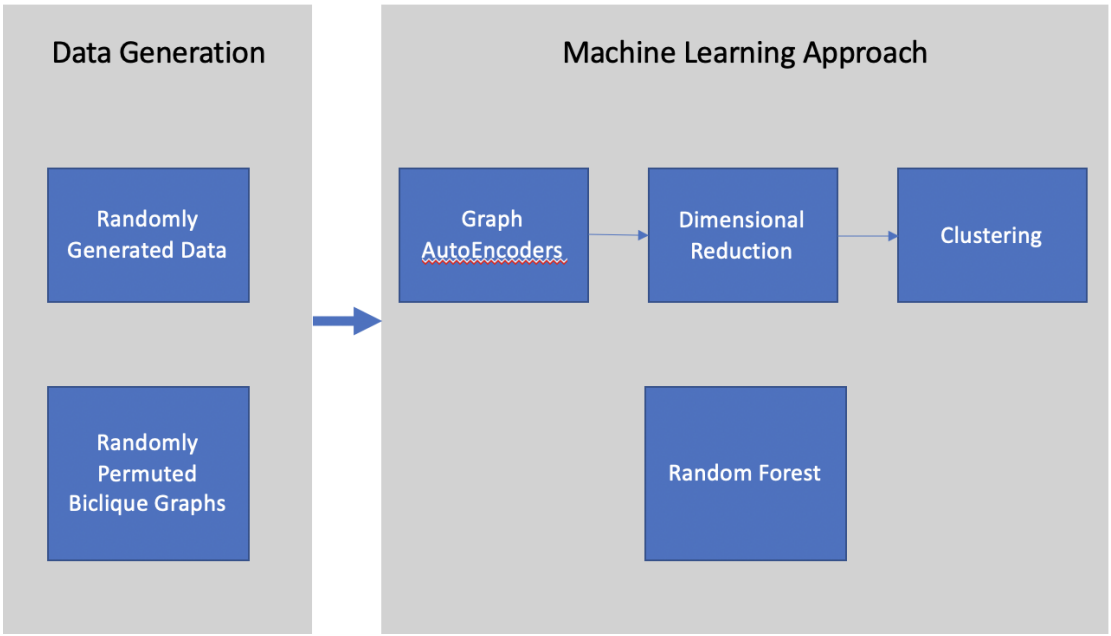


Figure 2.1: Load testing flow chart

Some generated HiSIM graphs proved to large for VZ plots to handle. Therefore, in order to load test the VZ visualizer, python programs were created. The pipeline used to load test the VZ plots can be seen in Figure 2.1. All code used to load test the VZ plots can be accessed on Github [2].

[2]https://github.com/shenoisam/Thesis

Data Collection

The first stage of load testing the VZ plots involved the generation of test data. Data was generated in two ways for this stage: Randomly Generated Dataset (RGD) and Randomly Permuted Biclique Graphs Dataset (RBG).

Data collected as part of the RGD dataset included 1725 biclique graphs that were randomly generated using a python program. The graphs were similar to those generated from the Geneweaver HiSIM graph, but were not based on existing genesets. Instead the graphs were created in order to ensure that a certain number of nodes, edges, and hierarchical depths were created. Graphs contained between 1 and 100 nodes (n), between 0 and n^2 edges (m), and up to n hierarchical depths. All 1725 graphs were tested using the VZ plots and manually annotated. The purpose of these graphs were to test the importance of n, m, and hierarchical depths in the ability for the VZ plots to display.

For the RBG dataset, we ran the Geneweaver HiSIM graph on randomly selected public genesets. We collected the successful runs(RBG-UnPermutated dataset), where the HiSIM Graph was able to run to completion, and tested the graphs to see if they could be displayed using the VZ visualizer. After annotating all of the collected graphs, we sampled 10 graphs from the RBG-UnPermutated dataset and randomly permuted them. The permutations removed all existing edges and redrew new random edges between nodes. During this permutation, the number of edges and the number of nodes were kept constant and the only difference between a permutated graph and the starting graph were the starting and ending nodes of the edges. One hundred permutations of the each of the 10 graphs were created using this method and were saved as a dataset (RBG dataset). Since we knew that these original graphs either displayed or didn't display, permutation allowed us to determine if any latent features irrespective of number of nodes or number of edges were important in the visualization of these graphs.

Machine Learning Approach 1

The next stage in load testing the VZ plots was to determine the relationship between the structure of the graphs and their representations in the VZ subway plots. To do this

we experimented with 2 different machine learning techniques - labeled as MLA1 and MLA2.

MLA1 sought to address the major challenge of accounting for latent features in the graph that would affect their visualizations. The existence of these features potentially prevented a simple comparison of number of nodes, number of edges, and node depth from being accurate.

In order to solve this problem, MLA1 tested two graph embedding techniques: GCN and linear. The purpose of both embedding techniques was to 1) preserve the structure of the original graph while 2) reducing the dimensionality of the graphs [111]. The Graph Convolutional Network (GCN) model works by iteratively aggregating the embeddings of neighbors for a node and then using a function of the obtained embedding and its embedding at previous iteration to obtain the new embedding [111, 112]. An extension of this model, Variational Graph Auto-Encoders (VGAE), uses a GCN encoder and an inner product decoder to encode the graph. This has been shown to improve performance [111, 113]. Furthermore, the linear model seeks to replace the GCN autoencoder used for the VGAE with a linear model . This model, while simpler, has been shown to achieve competitive results [114].

Using these models, we tested our graphs on 4 variations of these models: GCN Auto Encoder (gcn_ae), GCN Variational Auto Encoder (gcn_vae), Linear Auto Encoder (linear_ae), and Linear Variational Auto Encoder (linear_vae). Each model was trained using the RBG dataset. A python script was created to test each of these encoders on the obtained datasets. To encode the graphs we used the default parameters of the graph autoencoders. These parameters can be seen in Table 2.1.

After autoencoding the graphs, the graph representations were then reduced to a 1D representation in order to prepare for the clustering stage. The dimensional reduction was preformed using the Sklearn PCA package [115].

Finally, the 1D representations of the graph were then clustered using Sklearn KMeans[115]. The clustered graphs were then compared to their annotated classification to see if there was an underlying relationship between the structure of graphs that were displayed prop-

Parameter	Value
Learning Rate	.01
Epochs	200
Hidden1	32
Hidden2	16
Weight Decay	0
Dropout	0
Dimension	16

Table 2.1: Default parameters used for all models encoded.

erly by the VZ subway plots or not.

Machine Learning Approach 2

MLA2 made use of a random forest algorithm in order to determine which graphs were unable to be displayed by the Subway plots. MLA2 assumed that there were no significant inaccessible latent features inherent in the graph that would impede machine learning performance.

The random forest algorithm implementation used the Scikit-Learn package. [116]. Networkx was used to load in the graphs and extract the following features: n,m, maximum hierarchical depth, mode of hierarchical depths, graph density, and number of self loops [117].

The random forest algorithm was trained on RBG-UnPermutated dataset and tested on the RGD dataset. In order to test the effectiveness of each feature to the random forest learning process, we tried a variety of combinations of the 6 input features (Table 2.2).

2.3 Results

2.3.1 VZ Integration

The VZ subway plots were successfully integrated into Geneweaver as seen by the screenshot in Figure 2.2. The newly created python module was able to take the json output from the existing HiSIM tool and convert it to a format that could then be passed into to the VZ Subway plots. The current view visualization on Geneweaver was updated

Run Number	density	n	m	max depth	mode depth	self loops
0	■					
1	■	■				
2	■	■	■			
3	■	■	■	■		
4	■	■	■	■	■	
5	■	■	■	■	■	■
6		■				
7		■	■			
8		■	■	■		
9		■	■	■	■	
10		■	■	■	■	■
11			■			
12			■	■		
13			■	■	■	
14			■	■	■	■
15				■		
16				■	■	
17				■	■	■
18					■	
19					■	■
20						■

Table 2.2: Features used for each run of the Random Forest classifier.

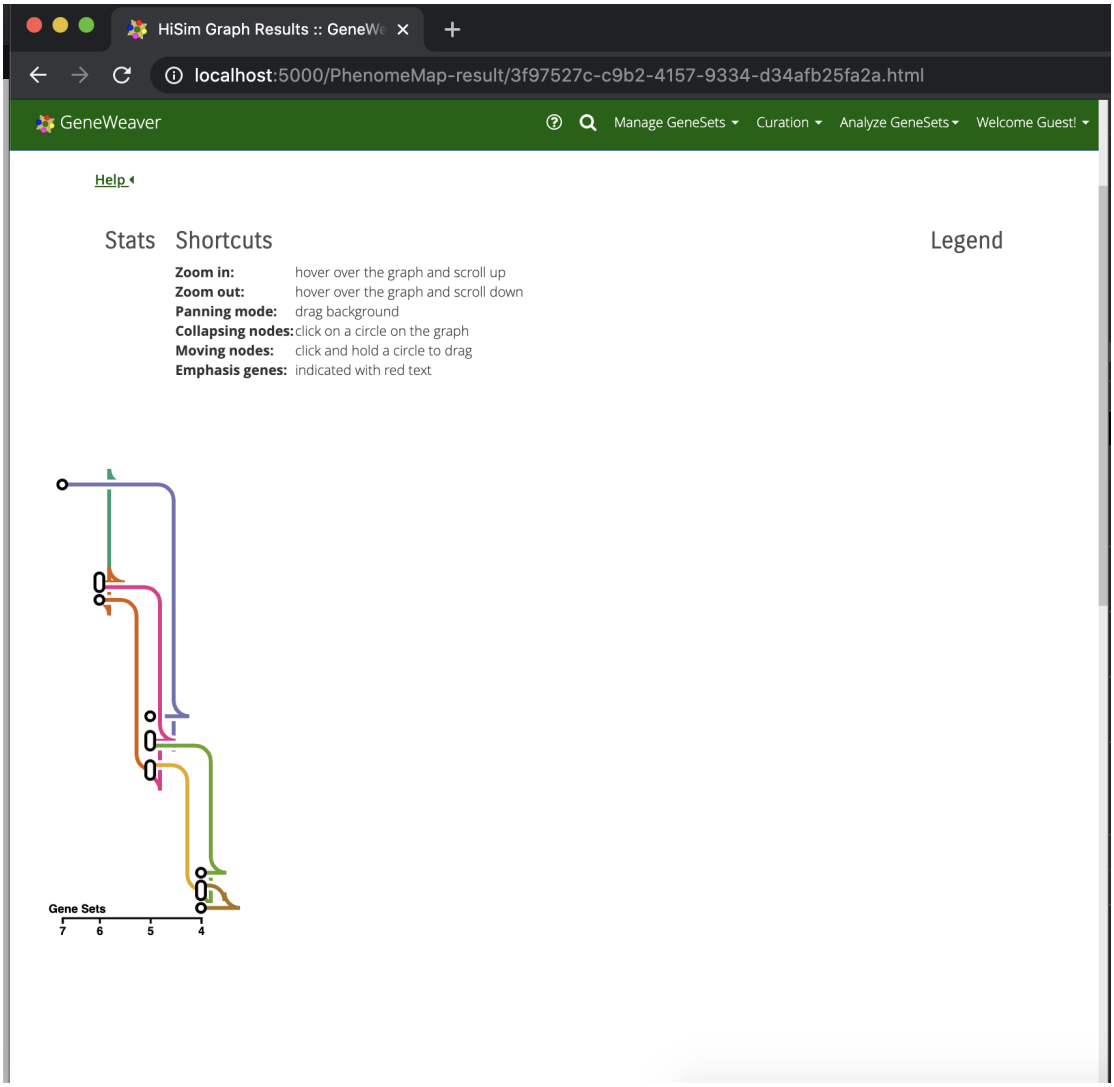


Figure 2.2: Integrated HiSIM Graph

Run Name	Number of Genesets	Num Permutations
RealRun0	19	100
RealRun4	2	9
RealRun6	10	100
RealRun7	4	36
RealRun102	11	102
RealRun105	10	105
RealRun16	4	16
RealRun108	10	108
RealRun117	12	100
RealRun124	3	25

Table 2.3: Data sets collected from running the HiSIM tool on Geneweaver with random genesets.

in order to allow users to switch between Subway plot visualizations and the existing HiSIM graph visualization.

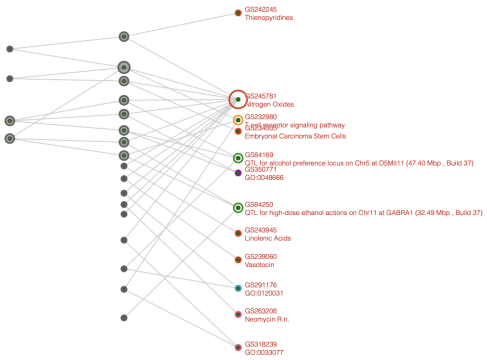
2.3.2 Load Testing

Data Collection

In order to construct the datasets required, automated data generation testing code ran the HiSIM graph 375 times. However, the majority of test cases failed due to a lack of genetic overlap between randomly picked genesets. Seventy eight successful trials were obtained and annotated (RBG-UnPermutated dataset). We then selected 10 graphs from this set to be used as initial starting graphs (Table 2.3, Figure 2.4). The selected initial starting graphs had between 2 and 19 randomly selected input genesets.

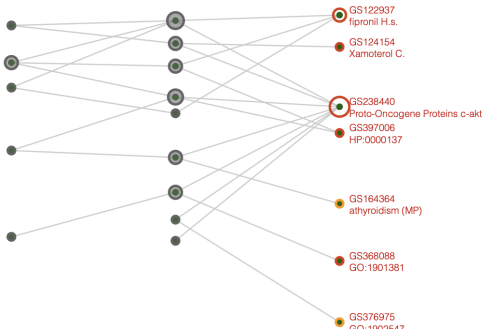
The 10 selected graphs were subsequently plotted using the VZ Subway plots as seen in Figure 2.6. Runs 105 and 16 graphical representations were not visible at all reaffirming original doubts that the VZ Subway plots could not display certain graphs. The remaining 8 graphs were able to be displayed.

These 10 resulting graphs were then permutated using a permutation python script. Runs 4,7,16, and 124 had 9, 36, 16, and 25 permutated runs respectively due to the lack of permutation options for such small graphs. The remaining 7 runs had the full 100 permutations. These combined permutations composed the RBG dataset.

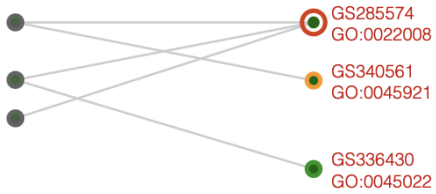


(a) RealRun0

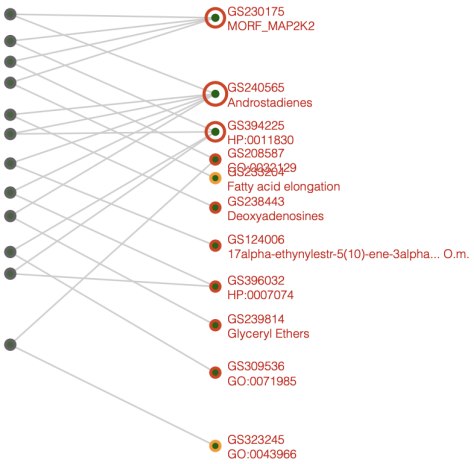
(b) RealRun4



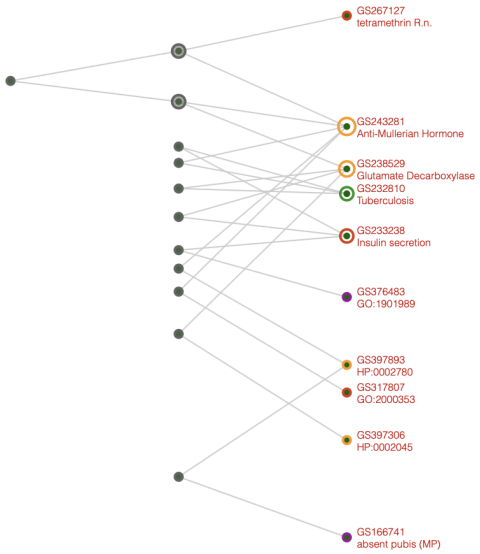
(c) RealRun6



(d) RealRun7

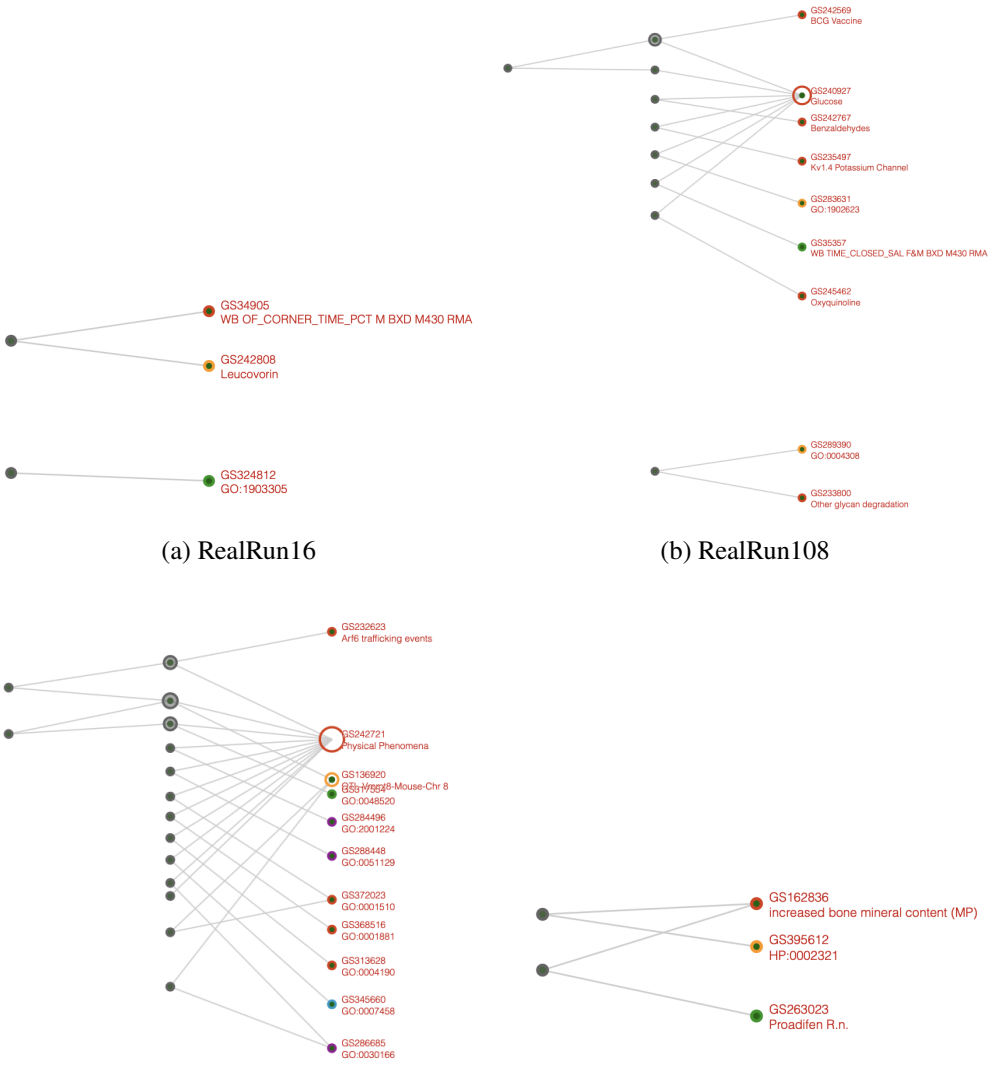


(e) RealRun102



(f) RealRun105

Figure 2.3: Random HiSIM graph runs



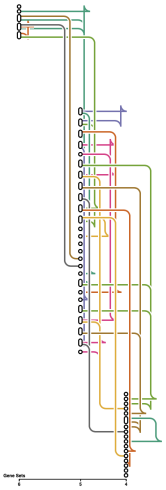
(a) RealRun16

(b) RealRun108

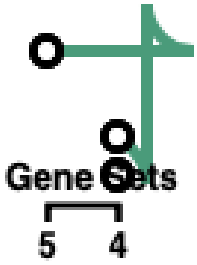
(c) RealRun117

(d) RealRun124

Figure 2.4: Random HiSIM graph runs continued



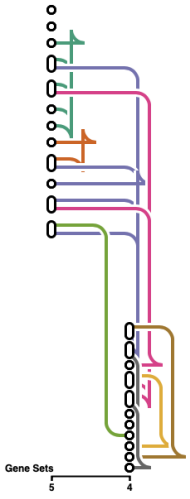
(a) RealRun0



(b) RealRun4

(c) RealRun6

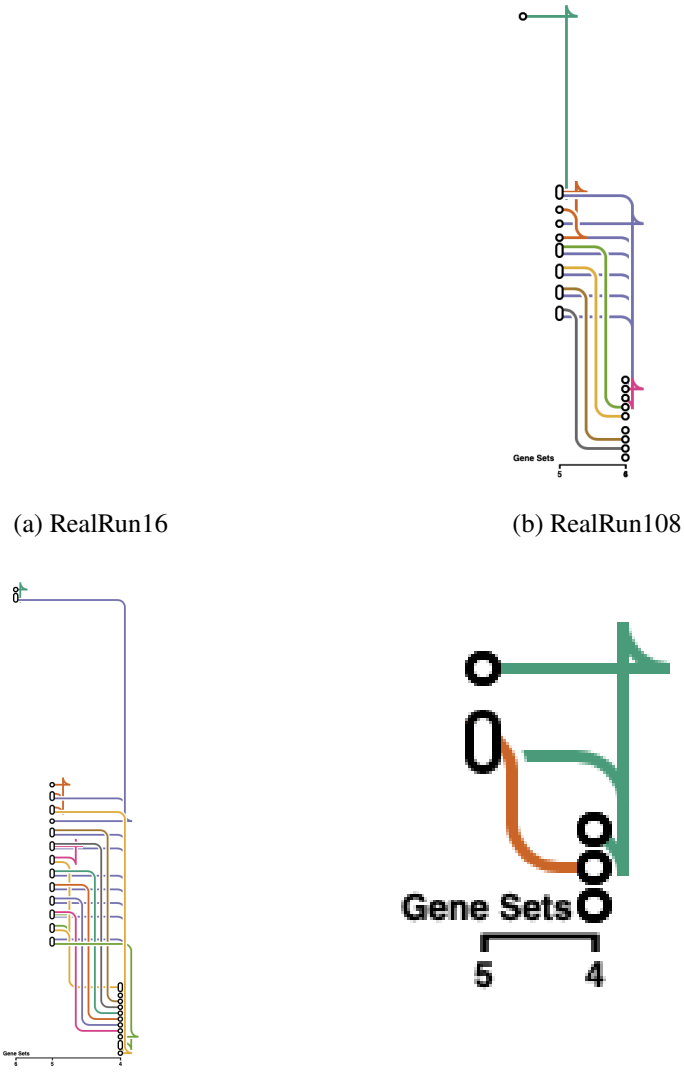
(d) RealRun7



(e) RealRun102

(f) RealRun105

Figure 2.5: Random HiSIM graph runs visualized using VZ Subway plots



(a) RealRun16

(b) RealRun108

(c) RealRun117

(d) RealRun124

Figure 2.6: Random HiSIM graph runs visualized using VZ Subway plots Cont.

Run	Encoder	AE Median ROC Score	AE Median AP Score	VAE Median ROC Score	VAE Median AP Score
RealRun0	gcn	.5	.621	.5	.633
RealRun0	linear	0.484	0.589	0.472	0.589
RealRun4	gcn	NA	NA	NA	NA
RealRun4	linear	NA	NA	NA	NA
RealRun6	gcn	NA	NA	NA	NA
RealRun6	linear	NA	NA	NA	NA
RealRun7	gcn	NA	NA	NA	NA
RealRun7	linear	NA	NA	NA	NA
RealRun102	gcn	.5	.5	.5	.667
RealRun102	linear	.5	.5	.5	.5
RealRun105	gcn	1	1	0	.5
RealRun105	linear	0	.5	0	.5
RealRun16	gcn	NA	NA	NA	NA
RealRun16	linear	NA	NA	NA	NA
RealRun108	gcn	0	.5	0	.5
RealRun108	linear	0	.5	0	.5
RealRun117	gcn	.5	.583	.5	.75
RealRun117	linear	.5	.75	.5	.583
RealRun124	gcn	NA	NA	NA	NA
RealRun124	linear	NA	NA	NA	NA

Table 2.4: Median Roc Score and AP score for each encoder tested on each of the datasets. Only 5 of the 10 tested datasets were able to be encoded.

Machine Learning Approach 1

The 10 graphs were then encoded using the 4 types of graph autoencoders: gcn_ae, gcn_vae, linear_ae, and linear_vae. The median ROC and AP score for each run can be seen in Table 2.4. Runs 4, 6, 7, 16, and 124 were unable to be encoded by any method. We believe that the size of these graphs was the major contributing factor to the inability to be encoded. Due to the poor ROC and AP scores that were obtained when encoding the graph, we decided not to proceed with the Dimensional Reduction and Clustering aspects of our pipeline. These low scores indicated that the encoders were unable to truly capture the representation of these testing graphs and therefore further processing using these embeddings would prove fruitless.

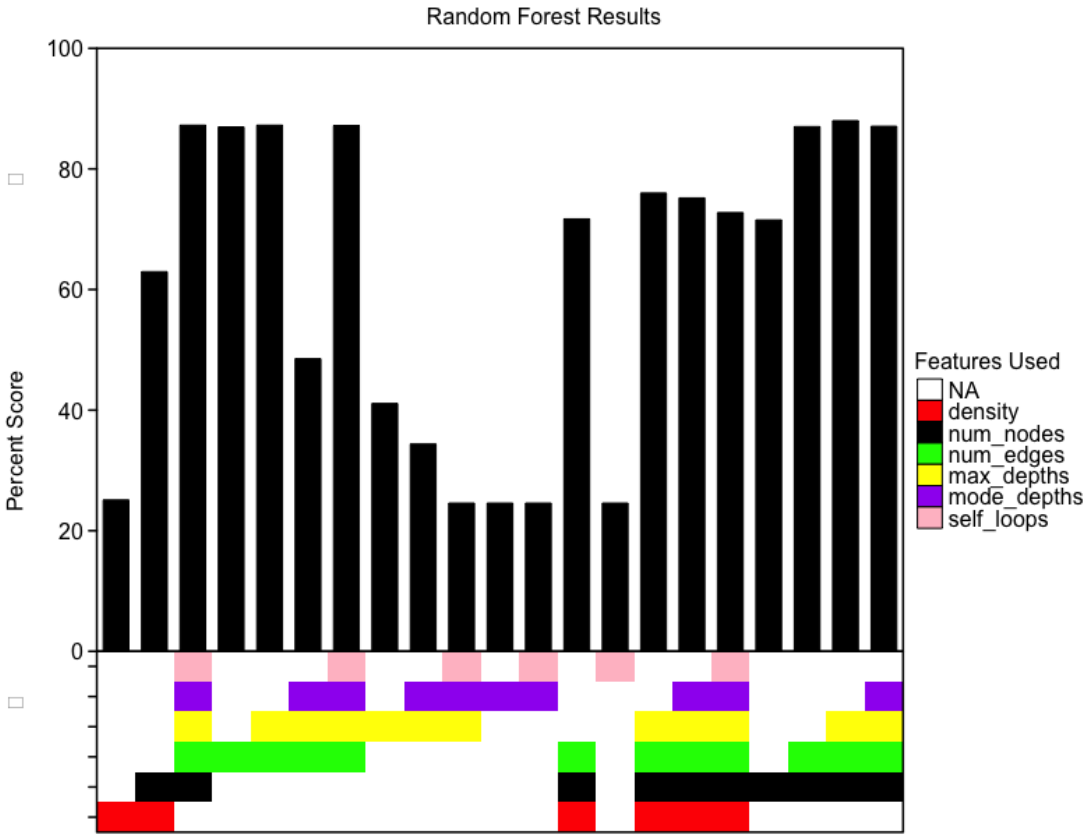


Figure 2.7: Results of the Random Forest run. Features were selected to handle the training and testing of the model. The best preforming run used only n,m, and the maximum depth. Graph was created using the Boutros Lab Plotting R package [118, 119]

Machine Learning Approach 2

The graphs were then tested using the MLA2 which utilized the random forest methodology. The classifier was trained on the RBG-UnPermutated dataset and tested on the RGD dataset. The results of each run and the features used to test the graphs can be seen in Figure 2.7. Using the Random Forest Methodology,the best run was able to correctly classify the graphs 88% of the time.

2.3.3 Prediction

In order to predict the node and edge values that the VZ plots could handle, we then employed the random forest model to classify varying node, edge, and depth levels. Figures 2.9 and 2.8 show the predicted number of nodes, edges, and maximum depths that the VZ plot could handle. Based on our analysis, the VZ plots are able to display

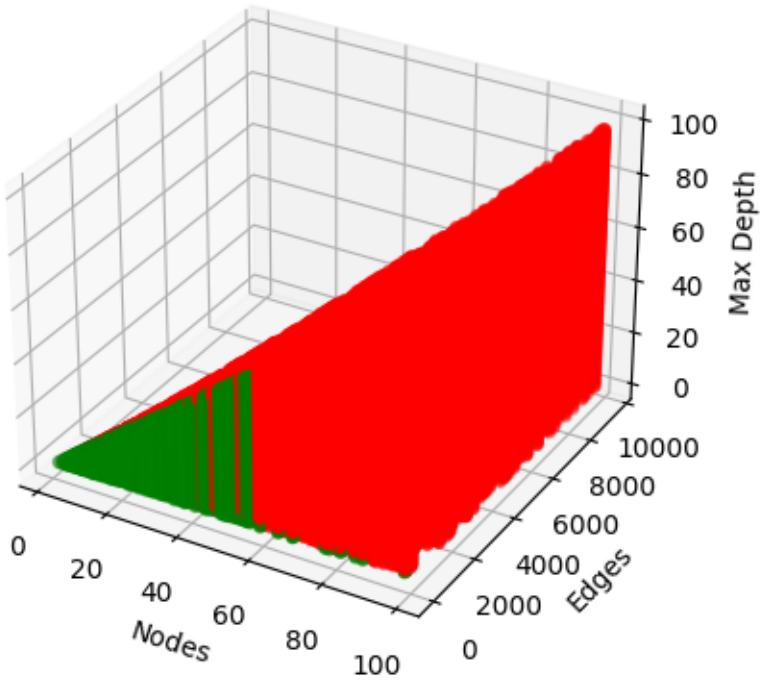


Figure 2.8: Number of nodes, edges, and maximum depths that the VZ plots can display.

graphs with around 50 n, 65 m, and a maximum of 30 hierarchical depths.

2.4 Discussion

Through this work, we were able to integrate a new visualization technique into the existing Geneweaver architecture. The added graphical implementation provides selective advantages over previous visualization techniques in that it provides a more colorful representation of the result graph and provides a key corresponding the number of genesets present per level.

We were able to determine that there exists a relationship between the number of node, edges, and maximum depths and the ability for the VZ plots to display correctly. This relationship is irrespective of any latent features of the graph. As n,m, and maximum depth increase, the VZ plots are less likely to be able to display the resulting graph.

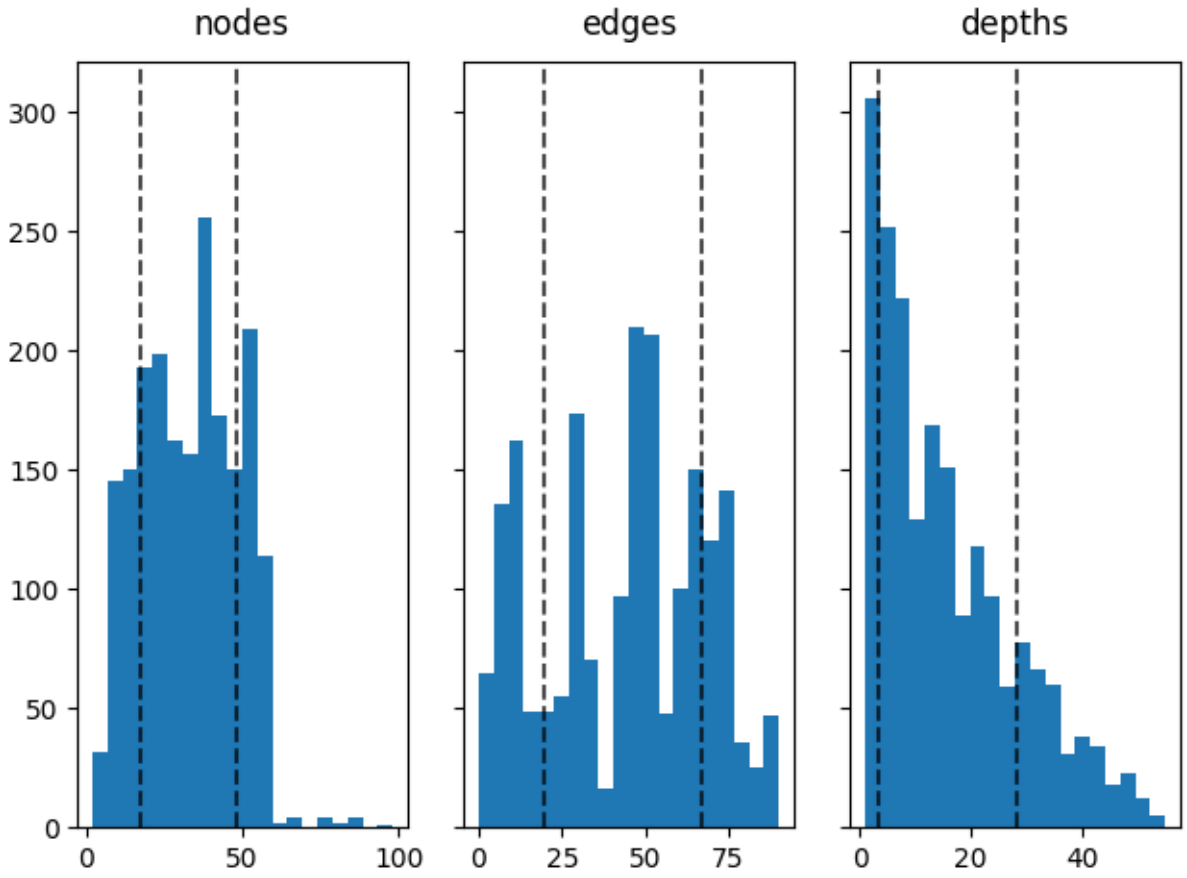


Figure 2.9: Histogram displaying the number of nodes, edges, and maximum depths the VZ plots can display. Dashed lines represent 1 standard deviation from the mean

An interesting finding that emerged from this study was the poor results obtained when using the GAEs. We initially tried changing the parameters used for encoding the graphs, but we were unable to see a significant improvement in our research (Data not shown). We believe that the sizes of the graphs and their unique structure were some of the contributing factors to the poor performance discovered.

CHAPTER 3

Network Analysis

3.1 Introduction

As previously mentioned in Chapter 1 of this thesis, the genetics of BD is of particular interest. In virtually every studied population, the HLA-B51 gene has been the most closely associated risk factor for BD [21]. However, the facts that sizable portion of BD patients present as HLA-B51 negative and the high prevalence of HLA-B51 in American Indian populations where BD is rare indicates the role of at least one other genetic loci as a factor of BD patheogenesis [21, 26]. Some potential candidates to fill this role have been identified including ERAP1, TNF-α, and IL-10 [12, 36, 33, 54, 10, 55, 56, 57]. The lack of a clear mechanism of BD patheogenesis and the genetic variation present in BD patients poses an interesting question: how can we separate the true genetic factors of the disease from the noise?

Furthermore, the classification of BD as an autoimmune disease and the discovery of isolated shared genetic factors between other autoimmune diseases and BD is also interesting. For instance, ERAP1 has been shown to be a susceptibility gene in BD, psoriasis, and ankylosing spondylitis while IL-10 has been associated with multiple other autoimmune diseases including rheumatoid arthritis, systemic sclerosis, Kawasaki disease, Sjogren's syndrome, Grave's disease, myasthenia gravis, psoriasis, autoimmune lymphoproliferative syndrome, and BD [10, 56, 58, 120]. This overlap between autoimmune diseases presents as significant area for research as gaining a deeper understanding of the shared genetic pathways of autoimmune diseases can have important implications for diagnosis, treatments, and future research[121].

In addiction research, where single genes rarely largely impact clinical phenotype and hundreds if not thousands of variants are needed to fully explain underlying genetics of the diseases, convergent functional genomics (CFG) is leveraged to make sense from

the noise. The premise of CFG is simple: the more lines of evidence for a gene, the higher it is prioritized as a gene of interest [121, 122]. This approach allows data from sources such as Genome Wide Association Studies (GWAS), gene expression studies, and animal models to be synthesized as evidence to ascertain the impact of genes on phenotype[122]. Even data from sources with limited sample size or lack of replication studies can be used as data points in the CFG approach since it is the collective evidence that together makes up the prioritization list.

To implement the CFG approach, we leverage the computational power of the Geneweaver HiSIM graph[8]. The hierarchical structure of the HiSIM graph output allows us to categorize genes based on where on the hierarchy they fall. Nodes at the top of the graph have a limited number of genes but many genesets as evidence while nodes at the bottom of the graph have many genes but fewer genesets as evidence. In this chapter, we use the CFG approach to first find candidate genes of interest for BD and then secondly determine other autoimmune diseases with genetic overlap with BD.

3.2 Methods

3.2.1 Data Collection

Genes associated with BD were collected as genesets on Geneweaver [7]. Each geneset contained a record of BD associated genes from a single study or source. Sixteen BD genesets were collected overall. Eleven of the genesets originated from GWAS studies and were collected by searching the GWAS Catalog and publicly available curated genesets from Geneweaver [123, 7]. The combined GWAS data came from a global population that included samples from Iranian, Japanese, Turkish, Korean, Spanish, Western European, Middle Eastern, and Han Chinese populations. [124, 125, 126, 65, 127, 128, 129, 62, 130, 65]. One of the GWAS studies collected data regarding BD and a special type of BD that effects the GI tract, Intestinal BD (IBD) [65]. This data source was split into two genesets for purposes of this study. Another geneset was created from the NCBI gene2mesh tool and was included in the study[131]. The Online Mendelian Inheritance

in Man (OMIM) database provided an additional geneset tagged "autoinflammatory, familial, Behçet's-like" [132]. Another geneset was created using information compiled in Malacards [133]. The last two genesets came from a BD review paper that consolidated genes of interest and a BD gene expression profile paper [33, 39]. Finally, we noticed that one of the existing BD genesets on Geneweaver did not include all of the genes identified by the GWAS study as associated with BD. To compensate for this, we created a new geneset from the same GWAS study and added all of the genes identified by the study as associated to BD. The union of all 16 genesets was subsequently collected using the Boolean Algebra tool on Geneweaver and stored as another geneset [7, 66].

Genesets for twenty-seven conditions were compiled in order to test BD 's relation to other autoimmune diseases. Disease genesets were created by using the Geneweaver Boolean Algebra tool on publicly available curated genesets [7]. Due to the large data size, only human genes were used for this run. The list of conditions and number of genes present in the genesets can be seen in Table 3.1.

3.2.2 BD HiSIM Run

To find the consensus genes of BD between all of the collected data sources, the Geneweaver HiSIM graph was run with the 16 BD genesets as input. The parameters used to run the HiSIM graph are reported in Table 3.2.

3.2.3 Autoimmune HiSIM Run

To determine the genetic overlap between BD and other autoimmune diseases. The HiSIM graph tool was run on 27 autoimmune disease genesets constructed using the Boolean Algebra tool and the BD Union geneset. The parameters used to run the HiSIM graph are in Table 3.2.

The data from the Autoimmune HiSIM run was then filtered to find relevant connections between BD and other autoimmune diseases. The resulting graphs were then plotted using the igraph python package [317] and the BoutrosLab.plotting.general R

package [119, 118]

3.2.4 Jaccard Geneset Analysis

After conducting the HiSIM run, a Jaccard Geneset analysis was conducted on autoimmune diseases identified as presenting with a high BD similarity. This analysis determines whether the geneset overlap identified in the Autoimmune Disease Run was statistically significant. The Jaccard Geneset analysis was run using the Jaccard Similarity tool on Geneweaver [7].

3.2.5 Neighbor Joining Analysis

In order to find the top five most genetically similar diseases to BD, a neighbor joining tree was created using the ape R package [318]. The distance between any two diseases was defined using the normalized Jaccard formula seen in Formula 3.1.

$$1 - \frac{A \cap B}{min(size(A), size(B))} \tag{3.1}$$

The formula took into account the differences in set size between the two sets in order to determine their overlap. By taking the size of the intersection between the two sets and dividing it by the largest possible overlap that could occur between the two sets, the formula allowed sets of different sizes to be compared. The distance between all possible pair combinations of the 16 genes was calculated and stored in a distance matrix. The distance matrix was then used as input to the neighbor joining program from the ape R package[318, 119].

3.3 Results

3.3.1 BD HiSIM Run

Running the HiSIM graph on the BD collected genesets resulted in the graph shown in Figure 3.1. Genes found in more than 3 genesets are displayed in Table 3.3. HLA-B

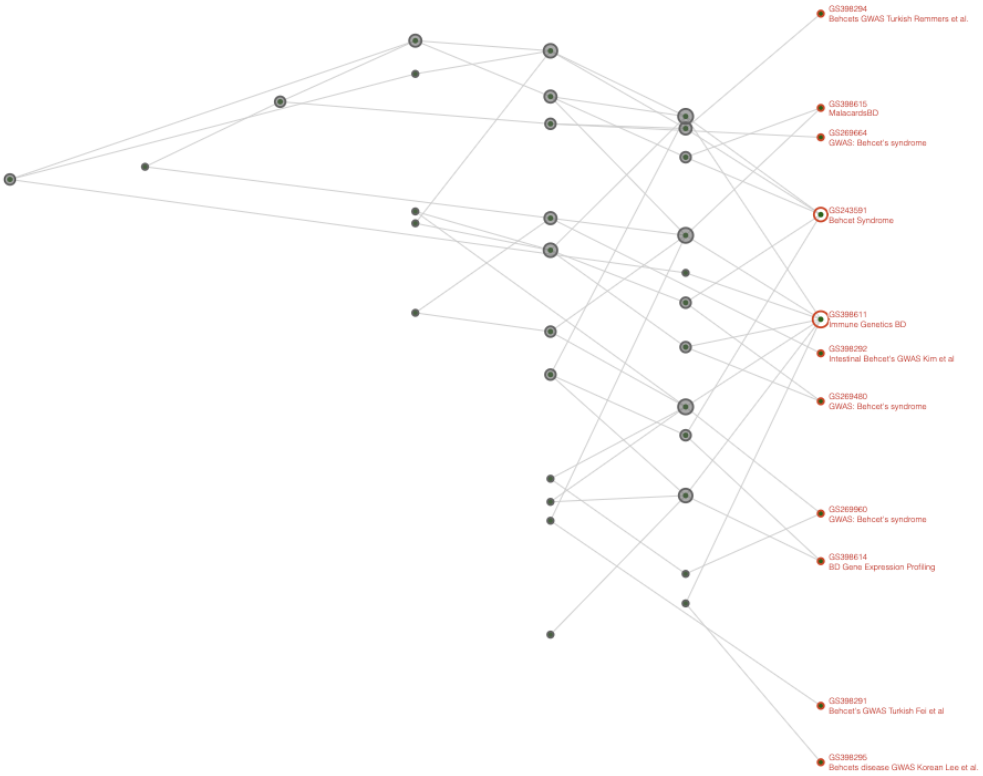


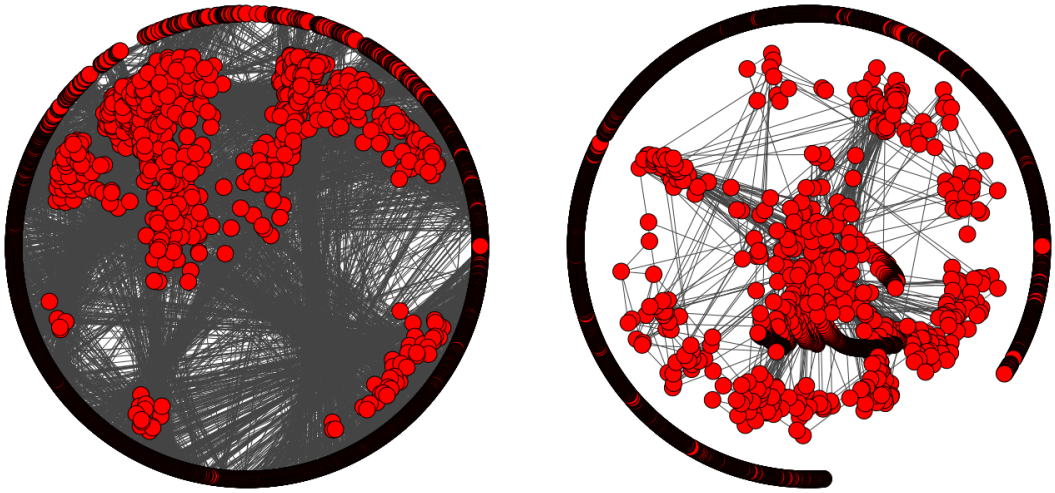
Figure 3.1: HiSIM Graph between 17 BD gensets.

was identified in 7 genesets making it the most common gene amongst all tested genes. IL-10 was the next most common gene and was found in 6 genesets. IL23R was the third most common gene, found in 5 genesets. Finally, HLA-A, STAT4, MICA, and ERAP1 were all found in 4 genesets.

3.3.2 Autoimmune Disease Run

After collecting autoimmune genesets on geneweaver, the HiSIM graph was run on 27 unique autoimmune diseases and BD. Figure 3.2 displays the results of this run. In Figure 3.2A, all edges between computed nodes are included. The resulting dataset was subsequently filtered in order to remove all nodes that are unrelated to BD; Figure 3.2B displays the results of filtering.

Figure 3.3 displays the genetic overlap between the autoimmune diseases and BD as determined by the HiSIM graph. Out of the 27 conditions tested, only 16 were found to have some genetic overlap with BD. The 16 identified candidates were then used to subsequently preform the Jaccard Analysis.



(a) Autoimmune HiSIM graph (b) BD Filted Autoimmune HiSIM Graph

Figure 3.2: Results from the Autoimmune HiSIM Graph run. 27 autoimmune diseases were compared to BD in order to find which diseases had the largest genetic overlap with BD. **A)** All edges between computed nodes were drawn. **B)** Edges were filtered such that only edges that connected to nodes containing the BD geneset were drawn.

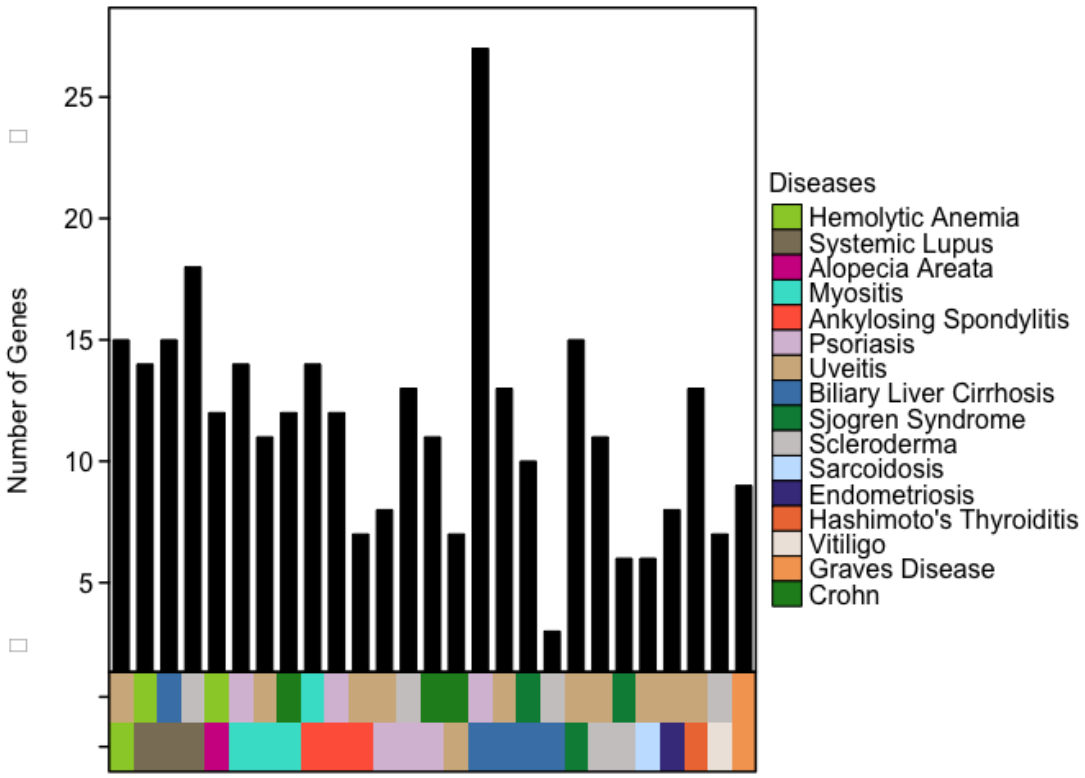


Figure 3.3: Graph displaying identified HiSIM nodes with the largest gene overlap with BD. 16 conditions were found to have some overlap with BD out of the 27 tested.

Figure 3.4: Results of the Jaccard Geneset Analysis. None of the 16 autoimmune disease genesets were identified as having a statistically significant overlap with BD.

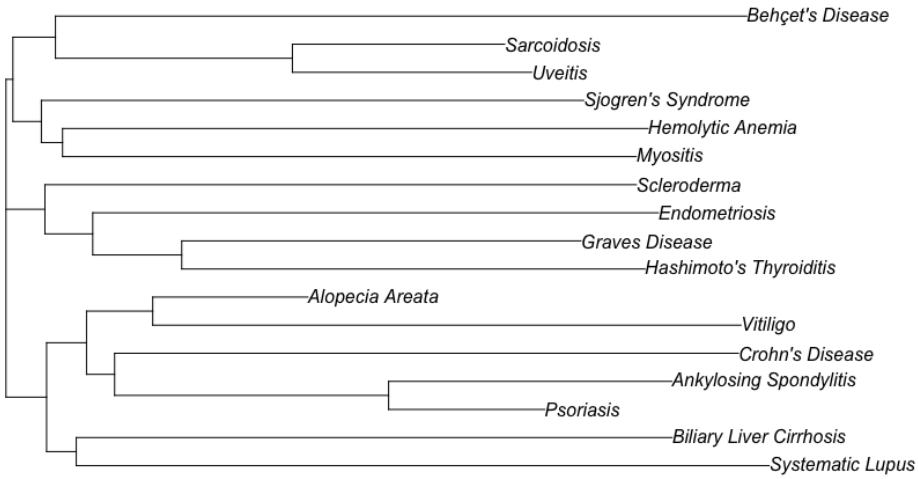


Figure 3.5: Results of the Neighbor Joining Tree.

3.3.3 Jaccard Geneset Analysis

Using the 16 identified conditions from the Autoimmune HiSIM run, a Jaccard Geneset Analysis was run. The results from the Jaccard Geneset analysis found that the overlap between all of the conditions and BD did not reach a level of statistical significance(Figure 3.4).

3.3.4 Neighbor Joining Analysis

The genetic overlap between BD and the 16 identified autoimmune diseases was then normalized using Formula 3.1. The normalized genetic overlap was then used as input for the Neighbor Joining Tree, the results of which can be seen in Figure 3.5. The tree indicates that BD is closest to Sarcoidosis, Uveitis, Sjogren's Syndrome, Hemolytic Anemia, and Myositis.

3.4 Discussion

The results of the BD HiSIM graph run are consistent with previous BD research. All top scoring genes have previously been identified as having an association with BD

[18, 33, 10]. The emergence of HLA-B as the top scoring gene is unsurprising given the history of HLA-B51 as a susceptibility gene of BD and its reputation as the the primary gene of interest in BD [21, 12, 18].

The results of the BD autoimmune disease run identified 16 diseases with at least some genetic overlap with BD. While these associations were unable to achieve statistical significance based on the results of our Jaccard Similarity run (Figure 3.4), we believe that these results can be potentially attributed to the limitations of the Jaccard Similarity test. It has been documented that the Jaccard coefficient, which is used to calculate the Jaccard Similarity, is strongly effected by dataset size [319]. This represents a limitation to our current study that can be addressed in future work.

Additional limitations to our study are primarily centered around data collection. Due to the limited genomics literature available for Behçet's Disease we were only able to collect 16 unique genesets for analysis. While sufficient for our current study, the addition of future GWAS and other genomic studies will help further elucidate the complex genomic interactions at play in BD. Similarly, our autoimmune disease HiSIM run, while large, is not exhaustive. All of the conditions collected were relatively common and data for these runs were limited to publicly available genesets already present in Geneweaver. This selection criteria excluded non-autoimmune conditions, such as cervical cancer, oral cavity cancer, and oropharyngeal cancer, that might mimic BD phenotype presentations.

To circumvent these limitations, we proposed a modified Jaccard similarity formula (Formula 3.1) to determine the genetic similarity between any two diseases. We then used this formula to construct a neighbor joining tree which allowed us to identify how genetically similar each of the 16 identified diseases were to BD.

The results of our CFG approach have the benefit of identifying autoimmune diseases with a shared genetic architecture, and is a promising strategy for identifying genes associated with BD. The inclusion of certain diseases in 16 identified diseases was expected. Most notably, Uveitis's high genetic overlap with BD was expected given that this condition is a symptom of BD [12]. Sjogren's Syndrome's inclusion in our list was expected

as well given that Sjogren's Syndrome often presents with other autoimmune diseases [320, 321]. Because of the multiple reports of Sjogren's Syndrome and BD, Sjogren's Syndromes association with BD has already been tested and no association between the two was found [322, 321].

Of interest was the close proximity of Sarcoidosis and BD in the neighbor joining tree (Figure 3.5). Historically, there have been very few reports of patients presenting with both BD and Sarcoidosis [323, 324], but there has been a recent uptick in BD patients presenting with both Sarcoidosis and BD after treatment with TNF antagonists [323]. Clinically, it can be difficult to distinguish between BD and Sarcoidosis considering the similarity of the articular, neurological, and skin lesions between the two diseases [324].

Additionally, the inclusion of psoriasis in our list was noteworthy as well. Psoriatic arthritis has been documented to sometimes be confused with BD articular involvement [325]. Hahn *et al.* recently found that psoriasis patients were twice as likely to be diagnosed with BD based on an analysis of patient data from the National Health Insurance Database of Korea [326]. Finally, Intestinal BD is sometimes misdiagnosed as Crohn's Disease and vice versa due to the similarity of symptoms, so its inclusion in our list of related diseases was unsurprising[327].

These findings, coupled with our genetic analysis, might provide evidence for a related underlying genetic mechanism related to the pathogenesis of these diseases. It will be left up to future work to fully clarify these relationships.

There have been documented case reports of patients presenting with both BD and one of the other diseases identified from our analysis. There have been a handful of cases of patients presenting with both BD and Sclerosis [328, 329], Myositis [330, 331, 332, 333], Endometriosis [334], Graves disease [335], Vitiligo [336, 337], Hemolytic Anemia [338], tuberculous thyroiditis [339, 340],and Biliary Liver Cirrhosis [341]. However, the inclusion of Alopecia Areata in this set of conditions was surprising. To our knowledge, there have been no reported cases of patients presenting with both Alopecia Areata and BD nor evidence of an association between BD and Alopecia Areata.

3.5 Conclusion

Behçet's disease is a complex multi-system inflammatory disease in which the exact pathogenesis continues to elude researchers [3, 1]. Genetically, the identification of HLA-B51 as a major, but not sole, susceptibility gene has led to the hunt for other genetic factors of the disease [21]. The resulting identification of multiple genes of interest, however, does not explicitly establish the contributions of each of these genes towards the overall presentation of the disease [12, 36, 33, 54, 10, 55, 56, 57]. Furthermore, research into these genes has uncovered roles of these genes in the progression of other autoimmune diseases [10, 56, 58, 120]. It is imperative to research autoimmune diseases in relation to other autoimmune diseases in order to fully understand human disease.

In this study, we employed a functional convergent genomics approach to discover 1) the genetic factors and 2) related autoimmune conditions of Behçet's Disease. The power of this approach lies in its ability to synthesize information from multiple genomics data sources [121] and in its recognition of shared genetic factors of autoimmune diseases [10, 55, 56]. The ability to quickly and accurately synthesize this information using the Geneweaver HiSIM graph presents a valuable opportunity for further discovery in this line of research [8]. Furthermore, our results using this approach confirmed existing BD research regarding the genetic factors of the disease, and identified 16 autoimmune diseases that share an underlying genetic relationship to BD. Almost all of these associations - the only exception being Alopecia Areata - have documented clinical findings linking them with BD, further providing evidence towards our results. It will be left up to further research to fully uncover the complex genetic interactions underlying these diseases as well as the shared genetic mechanisms between them.

Condition	Number of Genes	Number of Gene-sets	Publications
Achalasia	12	3	[134]
Vitiligo	107	10	[135, 136, 137, 138, 139, 140, 141]
Celiac Disease	428	22	[142, 143, 144, 145, 146, 147, 148, 149, 134, 150, 151, 152, 153, 154, 155]
Crohn's Disease	273	25	[156, 157, 158, 159, 147, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173]
Psoriasis	382	23	[174, 175, 176, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188]
Diabetes	130	20	[189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208]
Uveitis	130	7	[134]
Sarcoidosis	107	7	[209, 160, 210]
Myositis	82	5	[134, 211]
Scleroderma	148	9	[212, 213, 134, 214, 215]
Agammaglobulinemia	39	3	[134]
Alopecia Areata	26	3	[216]
Restless Leg Syndrome	27	6	[134, 217, 218, 219, 220]
Sjogren Syndrome	145	7	[221, 199, 222, 223]
Hashimoto's Thyroiditis	99	6	[134, 224, 225]
Addison Disease	40	4	[226]
Amyloidosis	58	12	[134]
Ankylosing Spondylitis	132	6	[227, 228, 150, 229]
Chagas Syndrome	159	7	[230]
MS	201	25	[231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254]
Endometriosis	226	9	[255, 256, 134, 257, 258, 259, 260]
Ulcerative Colitis	386	23	[261, 262, 263, 264, 265, 266, 160, 267, 163, 268, 269, 270, 271, 168, 272, 273, 274, 275, 276, 277, 278]
Hemolytic Anemia	261	20	[134, 279]
Systematic Lupus	269	25	[280, 281, 282, 283, 221, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 214, 295, 296, 297, 298, 299]
Liver Cirrhosis	130	8	[300, 301, 302, 134, 303, 304, 305]
Graves Disease	121	8	[306, 307, 308, 134, 309, 225, 310]
IgA nephropathy	39	6	[311, 312, 313, 314, 315, 316]

Table 3.1: Genetic information from 27 autoimmune diseases were collected from multiple publicly available genesets on Geneweaver. Genesets from each condition were then consolidated into one geneset using the Boolean Algebra tool.

Parameter	Value
Disable Bootstrap	False
Node Cutoff	Auto
Homology	Included
Genes In Node	All
Use FDR	False
Hide Unemphasized	False
p-value	1
Minimum Overlap	1%
Minimum Genes	1
Permutation Time Limit	5
Max In Node	4
Permutations	0
Max Level	40

Table 3.2: Parameters used for both runs of the Geneweaver HiSIM graph

Num Genesets	Gene	Genesets
7	HLA-B	Behçet's GWAS [65], BD Gene2Mesh, Behçet's GWAS [126], Behçet's GWAS [129], Malacards BD [133], Immune Genetics of BD [33], Behçet's GWAS [66]
6	IL-10	Intestinal Behçet's GWAS[65], Behçet's GWAS [66], BD Gene2Mesh, Behçet's GWAS [127], Malacards BD [133], Immune Genetics of BD [33]
5	IL23R	BD Gene2Mesh, Behçet's GWAS [66], Malacards BD [133], Immune Genetics of BD [33], Behçet's GWAS [127]
4	HLA-A	BD Gene2Mesh, Behçet's GWAS [126], Behçet's GWAS [66], Immune Genetics of BD [33]
4	STAT4	Immune Genetics of BD [33], BD Gene2Mesh, Behçet's GWAS [62], Behçet's GWAS [128]
4	ERAP1	Intestinal Behçet's GWAS[65], Malacards BD [133], Behçet's GWAS [128], Immune Genetics of BD [33]
4	MICA	BD Gene2Mesh, Behcet's GWAS [62], Immune Genetics of BD [33], Behçet's GWAS [66]

Table 3.3: Abbreviated results from the BD HiSIM run. IL10 was identified in 7 genesets making it the most common gene amongst all tested genes. HLA-B was the next most common gene and was found in 6 genesets. IL23R was the third most common gene; it was found in 5 genesets. Finally, HLA-A, STAT4, MICA, and ERAP1 were all found in 4 genesets.

CHAPTER 4

Autoimmune Disease Comparison

After identifying autoimmune diseases that were genetically similar to BD, the final stage of this thesis was to identify genetically and phenotypically representative animal models of BD. Since autoimmune diseases with a genetic relationship to BD were discovered previously in Chapter 3 (Figure 3.5) [342], the animal models for these conditions provided a starting point from which representative BD animal models could be identified from. These models could then be subjected to the International Criteria for Behçet's Disease (ICBD) in order to determine whether the phenotypic presentation of these models mirrored BD[16]. In this chapter, the discovery of novel animal models for Behçet's Disease is described.

4.1 Methods

4.1.1 Data Collection

The top 5 genetically similar diseases identified previously - uveitis, myositis, Sjögren's syndrome, sarcoidosis, and hemolytic anemia - were taken to be studied [342]. For each of the diseases, a brief review was conducted to identify clinical manifestations and current animal models.

Animal models were identified by researching literature and by using The Jackson Laboratory Human - Mouse: Disease Connection website [343]. For each animal model identified, the literature describing the animal model was scanned in order to find out if the model presented with any potential symptoms of BD. For the purposes of this study, potential symptoms of BD included ocular, oral, skin, genital, neural, gastrointestinal, or vascular involvement.

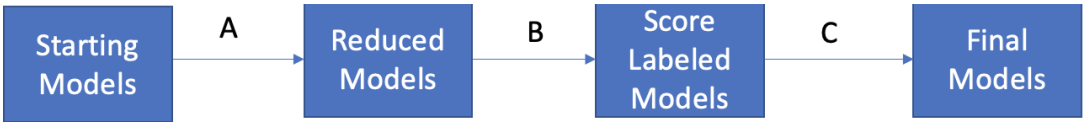


Figure 4.1: Workflow describing the evaluation of animal models. **A.** Remove all models with no BD areas effected. **B.** Score phenotypes based on BD diagnostic criteria **C.** Assess genetic relevance of models

4.1.2 Mouse Model Analysis

After compiling the list of animal models for each of the 5 diseases, the compiled models were evaluated to determine relevance (if any) to BD using the workflow defined in Figure 4.1. The first step (Step A) removed models with no apparent BD symptoms based on the results from the Data Collection stage. This resulted in a group of animal models termed as Reduced Animal Models.

The remaining models were then subjected to the International Criteria for BD (ICBD) and were thus assigned a numerical score (Step B) resulting in a group termed Score Labeled Models. The ICBD criteria uses a point based scoring system and a score of 4 is needed to be diagnosed with BD[16]. Based on the ICBD criteria, oral aphthosis, genital aphthosis, ocular lesions (anterior uveitis, posterior uveitis, or retinal vasculitis each add 2 points to the patients score. Neurological manifestations, skin lesions (pseudofolliculitis, skin aphthosis, erythema nodosum) and vascular manifestations (arterial thrombosis, large vein thrombosis, phlebitis or superficial phlebitis) each add 1 point to the patients score. For the purposes of this study we defined oral aphthosis as follows: a well-defined painful round or oval ulceration with a white base surrounded by a red areola that are localized to the lips, buccal mucosa, soft palate, and tongue of patients [3, 1, 11]. Similarly, we defined genital aphthosis as follows: a well-defined painful round or oval ulceration with a white base surrounded by a red areola most commonly found on the scrotum of men and the vulva in women but can also be found on the epididymis and penis of men and vagina and cervix in women [3]. Scoring the models in this way allowed for an easy comparison between models and the point values for each symptom can be seen in Table 4.1.

The first two steps primarily looked at phenotypic presentation of the models, so the

Symptom	Point Value
Oral aphthosis	2
Genital aphthosis	2
Ocular lesions	2
Skin lesions	1
Neurological manifestations	1
Vascular manifestations	1

Table 4.1: Points associated with each BD symptom based on the ICBD criteria

next step (Step C) assessed the genetic relevance of each model resulting the final group of models appropriately termed Final Models.

4.2 Results

4.2.1 Data Collection

The first disease to be researched was uveitis. Anterior and posterior uveitis have been shown to be symptoms of BD which made it an interesting target for further research [12]. Uveitis defines a group of diseases that present with intraocular inflammation involving the iris, ciliary body, and choroid. It can be caused by a localized infection or by a non infectious etiology [344]. Localized infections that have been shown to cause uveitis include Herpes Simplex Virus (HSV), varicella virus, cytomegalovirus, syphilis, toxoplasmosis, and tuberculosis [345]. On the other hand, non infectious uveitis has been shown in diseases such as BD, sarcoidosis, Vogt–Koyanagi–Harada (VKH) disease, juvenile rheumatoid arthritis, and multiple sclerosis [344]. There is some evidence to suggest that non-infectious uveitis may be initiated by an infectious agent as well, but the specific agent or mechanism of action has not been identified so far [346]. Thirteen animal models for Uveitis were identified from review papers and the Mouse Genome Informatics Website [347, 343]. All of the animal models identified presented with ocular lesions [348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361]. Two models, both using a version of the AIRE infection protocol, presented with lymphocyte infiltrates in retina, salivary glands, and ovaries [357]. One of the AIRE infection protocol mouse models also presented with lymphocyte infiltrates in the stomach [357].

The animal models associated with uveitis were compiled and listed in Appendix A.

The second disease to be researched was Myositis. Myositis, or Inflammatory myopathies, refers to a broad range of myopathies that affect not only the muscles but multiple organs. The forms of myositis include dermatomyositis, Overlap myositis, Necrotizing myopathy, polymyositis, and inclusion body myositis [362]. The animal models associated with Myositis were compiled and listed in Appendix B. Eighteen models in total were identified in mice, C. elegans, dogs, rats, guinea pigs,and hamsters[363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374]. Of the eighteen models identified, four had phenotypic presentation in areas of interest : Transgenic Mice V, Transgenic Mice VI, Infectious Mouse II, and Immunological IV [368, 369, 371]. Transgenic Mice V presented with skin collagen increase, transgenic mice VI presented with rashes on the skin and eyelid, Infectious Mouse II presented with inflammatory cell arthritis and multifocal vasculitis, and immunological IV presented with inflammatory cell arthritis.

The third disease to be researched was Sjögren's Syndrome. Sjögren's syndrome (SS) is an autoimmune disorder caused by the lymphocytic infiltration of exocrine glands resulting in glandular dysfunction, preferentially of the salivary and lachrymal glands [320]. The disease usually results in patients having ocular and oral dryness accompanied with salivary gland swelling. Additionally, patients frequently present with skin, nasal, and vaginal dryness, and between 30% to 50% of patients with the disease may present with articular, pulmonary, neurological, or kidney involvement [375]. The locations of Sjögren's syndrome phenotypic presentation, coupled with the discovered genetic overlap with BD, makes it an interesting target for further research. Specifically, the locations of the main symptoms of Sjögren's syndrome - eyes, mouth, skin, and vagina - mirror the locations of the main BD symptoms. The animal models associated with Sjögren's syndrome were complied and listed in Table Appendix D. Twenty two mice models were identified in total [376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399]. All of the models of Sjögren's Syndrome had either salivary or lacrimal gland involvement or both. Four models presented with additional symptoms: Mouse Model Spontaneous I

presented with conjunctivitis, Mouse Model Spontaneous II presented with testicular lesions, Sjögren's AIRE KO presented with stomach lymphoid cell infiltration, and Mouse T cell specific P13K KO presented with intestinal infiltration [377, 378, 379, 387, 388].

The fourth disease to be researched was Sarcoidosis. Sarcoidosis primarily presents with bilateral hilar lymphadenopathy and reticular opacities in the lungs and can also be found in skin, eyes, and joints. Additionally, it can express in the musculoskeletal system, reticuloendothelial system, exocrine glands, heart, kidney, and central nervous system [400]. Similarly to the other diseases researched, the genetic overlap between BD and Sarcoidosis combined with the phenotypic genes of interest made it an interesting target for further research. The animal models associated with Sarcoidosis were compiled and listed in Appendix C.Ten unique models were identified in either mice or rats [401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421]. Of the ten, only three presented with symptoms in at least one BD affected area: ApoE-deficient Mice, Tsc2 Mice, and Horse Sarcoidosis [408, 414, 415, 416, 417, 418]. ApoE-deficient Mice and Horse Sarcoidosis both presented with skin lesions and extra collagen in the GI tract. Tsc2 mice presented only with skin lesions. Interestingly, despite the fact that human Sarcoidosis effects the eyes, none of the models identified presented with any ocular manifestations [400].

The final disease to be researched was Hemolytic Anemia. Autoimmune Hemolytic Anemia is a decompensated acquired hemolysis caused by the host's immune system acting against its own red cell antigens[422]. Patients with this disease can remain asymptomatic and might only present symptoms when erythrocytosis cannot match the pace of red cell destruction. Other manifestations of this disease include jaundice, cholelithiasis, or isolated reticulocytosis. Patients may acquire Hemolytic Anemia through a variety of means that fall either in the acquired or hereditary category[423]. The animal models associated with Hemolytic Anemia were compiled and listed in Appendix E. Thirty seven models of the disease were discovered in mice[424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 456]. Of the models identified, two - ARTX KO

mice II and nb/nb mice - contained at least one BD symptom. ARTX KO II contained decreased horizontal and amacrine cells in retina while nb/nb mice contained neurological symptoms of an unbalanced gait and moderate tremor, indicating a nervous system involvement.

4.2.2 Animal Models Analysis

The identified models were then filtered to remove models that didn't affect one area present in BD (Step A in Figure 4.1). This resulted in 13 uveitis, 4 myositis, 22 Sjögren's syndrome, 3 sarcoidosis, and 2 hemolytic anemia models. Since both Uveitis and Sjögren's syndrome make use of the AIRE model, the total number of unique potential animal models remaining after Step A was 43.

Diagnostic Criteria Results

These 43 models were then subjected to the diagnostic criteria for Behcet's disease. In order to be classified as having BD, the models needed to score at least a 4 on the International Criteria for BD. The resulting scores of each model can be seen in Figure 4.2 and Appendix F.

The six highest scoring models scored a three on the ICBD criteria. These models all presented with vascular manifestations in addition to either ocular or genital lesions. These top scoring models were models for either uveitis or Sjögren's syndrome. Eleven models scored a 2 on the ICBD criteria, and all of these models with uveitis symptoms. These models were models for uveitis, Sjögren's Syndrome, and hemolytic anemia.

The remaining models tested presented with phenotypic presentation not inline with ICBD criteria despite phenotypic presentation in an organ or area of interest. For example, every Sjögren's syndrome animal model identified presented with either salivary gland or lacrimal gland infiltration. However, even though that lesions in the oral cavity and the eye are symptoms of BD, these specific symptoms receive a score of 0 on the ICBD criteria since they do not present in the correct area of the mouth and eyes respectively.

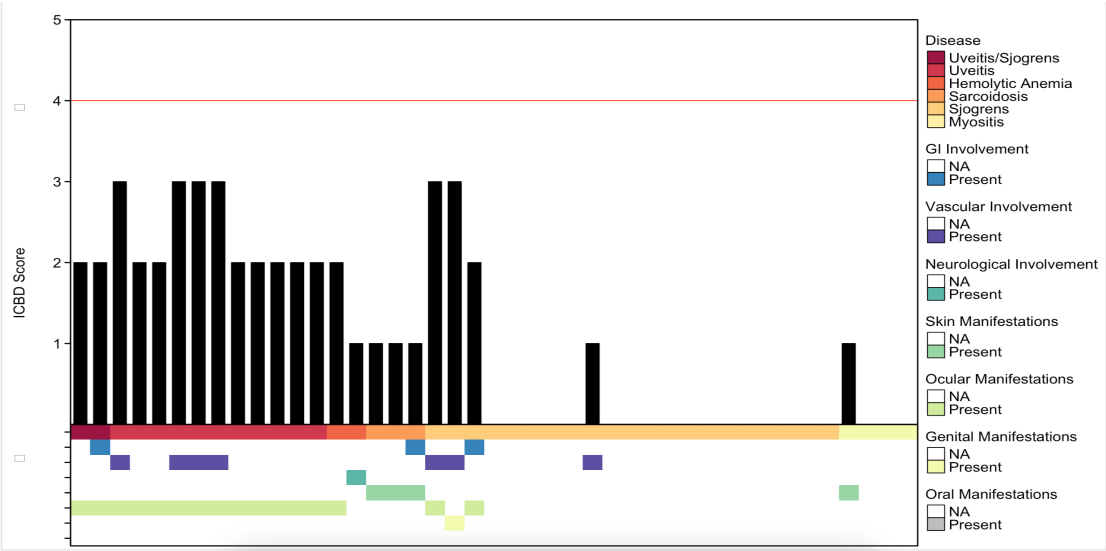


Figure 4.2: Scores of each animal model using the ICBD criteria. Models need to score a 4 or greater in order to be classified as having BD.

Genetic Analysis

Since no models passed the phenotypic test, the highest scoring models were then run through a genetic analysis in order to determine how genetically similar they were to BD. Seventeen models in total scored a 2 or greater on the ICBD criteria and, as such, were subjected to the genetic test. The results of the analysis can be found in Table 4.2.

Model	Gene	BD Gene
AIRE KO	AIRE	No
AIRE infected	AIRE	No
Transgenic HLA-A29	HLA-A29	Yes [33]
Induced Uveitis	NA (ocular antigens)	NA
Uveitis Humanized Tg	HLA-DR3	No
RPE65 Rat	RPE65	No
DC-EAU	IL-1, IL-6, IL-12, TNF-α, IFN-γ, and IL-10	Yes[67, 65, 66]
EMIU, EAAU	NA	NA
EAU bovine	NA	NA
EIU	NA	NA
Uveitis Double Tg	TCR α and TCR β	No
Experimental CMV retinitis	NA	NA
Expermimental animal ocular tuberculosis	NA	NA
ARTX KO mice II	ATRX	No
Sjogren's Mouse Spontaneous I	lpr	No
Sjogren's Mouse Spontaneous II	NA (NOD mice)	NA
Mouse T cell specific P13K KO	P13K pathway	Yes[39]

Table 4.2: Genetic Analysis of top scoring models.

Of the seventeen models subjected to the genetic test, seven were deemed ineligible due to the lack of a genetic component in the infection technique used to generate the model. Of the remaining 10 models, 3 were determined to have a genetic relationship to BD and 7 were determined to have no relation to BD.

4.3 Discussion

Our results indicate that 3 animal models - HLA-A29, DC-EAU, and P13K KO - might prove to be promising models for BD research.

The HLA-A29 emerged as one of the most promising models identified since it was one of the highest scorers on the ICBD criteria and presented with a genetic relationship to BD[34]. Phenotypically, the model presented with uveitis symptoms and vascular manifestations. Genetically, a relationship between HLA-A29 and BD has already been previously identified. The relationship showed increased HLA-A29 in patients presenting with ocular symptoms when coupled with a decrease in HLA-B51 [34]. Additionally, HLA-A29 has been shown to work in conjunction with variants in ERAP1, another implicated gene in BD patheogenesis, to induce birdshot uveitis [54, 10, 457, 458].These findings may prove to be interesting since ERAP1's role in BD patheogenesis revolves around it's interactions with HLA-B51. Therefore, the findings that ERAP1's interactions with HLA-A29 can also induce uveitis symptoms may provide an area for future research for HLA-B51 negative BD patients [54, 10, 457].The HLA-A29 mouse model may prove to be an effective tool in modeling HLA-B51 negative BD the future.

The Dendritic Cell Experimental Autoimmune Uveitis (DC-EAU) model was another one of the top scorers in our analysis since it presented with a phenotypic and genotypic component. The DC-EAU model scored a 3 on the ICBD criteria and presented with expression of multiple genes of interest. Classification wise, it falls in the category of an induced animal model of uveitis. Induced models of uveitis have been used in BD research before, and the most notable ones prescribed for BD research include interphotoreceptor retinoid-binding protein (IRBP) in BIO.A, BlO.S (8R), A, and

MRL (+/+) mice, recoverin in Lewis rats, and retinal soluble antigen or S-Ag in SJL/J X AKR, and AKR mice [6, 84]. The DC-EAU model's expression of IL-10, IL-6, TNF-α, IFN-γ and IL-12 in conjunction to ocular manifestations made it a promising model for BD research since these members of the interleukin family have been repeatedly shown to be associated with BD [67, 65, 66].

Finally, the mouse T-cell specific PI3K KO mouse presented with ocular lesions and included a BD gene of interest[39, 376, 388]. PI3Ks are a family of lipid kinases of which the the tumor suppressor gene PIK3R1 is a member[459]. The PIK3R1 gene was identified as having an association with BD in a study that used a gene array strategy to identify transcriptional profiles of peripheral blood cells obtained from patients with active BD.[39]. The PI3K signaling pathway is involved in a broad range of cellular processes, including growth, metabolism, differentiation, proliferation, motility, and survival, and class IA PI3Ks, of which PIK3R1 is a part of, stand out for their critical role in the mammalian immune system [460]. Mutations in the pathway have been classically associated with cancer, but recent research has indicated that mutations in the PI3K have led to immune dysregulatory disease with autoimmune manifestations and immunodeficiency[461, 460, 459, 462]. This model represents an interesting area for further research to determine the role, if any, the PI3K pathway plays in BD pathogenesis.

While our results didn't find a evidence of a genetic relationship between the AIRE models and BD, these models are worth mentioning here due to the facts that they are used as models for both Sjögren's syndrome and Uveitis and present with symptoms in the eyes, mouth, and genitals, all areas important in BD diagnosis. The AIRE models scored a 2 on the ICBD criteria due to the presence of ocular lesions as the oral and genital lesions were not localized in the correct area. Of interest is recent research indicating a possible diminishing of ovarian reserves in BD patients as well as a possible higher resistance and lower blood flow in the ovaries of BD patients [463, 464]. These results, while not conclusive, may provide evidence of an ovarian involvement, lending evidence towards the potential use of the AIRE model as a model of BD. Further

research is needed to confirm these results.

The first potential source of error in this study stems primarily from data collection. Since not all manuscripts document phenotypes unrelated to the condition that the model is targeted for, models presenting with BD symptoms might have been overlooked.

Despite the genetic and phenotypic presentation of the models discovered, the perfect animal model for BD remains elusive. The genetic relationships associated with 2 of the models recommended by this study were not identified as part of the convergent functional genomics approach outlined in our earlier work [342]. Further research is needed to explore these genetic relationships and their impact on BD patheogenesis. It is worth reiterating the potential of creating a transgenic HLA-B51 mouse model infected with HSV as mentioned by Charles *et al.* and Islam et al. [87, 6]. However, at present, this model has yet to be tested.

CONCLUSION

Behçet's disease (BD) is a multi-system inflammatory disease that affects patients along the historical silk road. The complex, potentially syndromic, nature of BD has made the identification of animal models for the disease quite challenging [6, 2]. Current animal models of the disease have been unable to replicate both the phenotypic and genotypic aspects of the disease [6] which proves to be a hindrance to the discovery of improved treatments for the disease. In this work, we undertook a computational approach to discover an improved animal for the BD that is able to encompass both the genotypic and phenotypic aspects of the disease.

To start, we first conducted a literature review that looked at the phenotypic and genetic factors of BD (Chapter 1). These factors where then used to analyze the current BD animal model landscape. The complicated phenotypic presentation and the existence of multiple BD subtypes makes developing BD animal model a major challenge as we discovered in our review. BD disease manifestations are present in multiple areas and as such obtaining a model that is able to replicate all clinical findings is challenging. Additionally, the proposed syndromic nature of BD [2] further complicates animal model creation since there exists a possibly that an animal model for one BD subset will be ungeneralizable to other BD subsets. On the genetics side, the lack of a clear mechanism of action severely limits the ability of models to replicate the full extent of the genetic interplay present in BD. Of note was the association between HLA-B51 and BD. In virtually every population studied, HLA-B51 has been the most closely associated risk factor for BD, but the incomplete penetrance of the gene and the high prevalence of HLA-B51 in American Indian populations where BD is rare indicates the role of at least one other genetic loci as factors of BD patheogenesis [21]. Finally, analysis of the current animal model landscape of BD revealed similar difficulties with the creation of an animal model. To date, the most widely used model for the disease is the infection of mice with Herpes Simplex Virus [85, 88]. This model is able to present with all of the same

symptoms as BD, but the lack of genetic influences in the model and the 33% mortality rate of the mice tested makes this model only feasible for large scale studies[6, 85]. A genetics approach towards developing a BD model also was partially successful as the HLA-B51 mice phenotypically presented only with an increased neutrophil count [89]. Thus, the need for the identification of an improved animal model is great.

To address this issue, we employed the Geneweaver HiSIM graph which allowed us to view a hierachial network of geneset interactions[8]. In order to improve the visualization of the graph, we integrated an improved visualization library (VZ plots) and load tested the plots in order to determine what types of graphs the plots could handle (Chapter 2).

After integrating the VZ plots into the Geneweaver HiSIM graph, we then used the graph to conduct a Convergent Functional Genomics (CFG) network analysis in order to find consensus genes and related autoimmune diseases to BD (Chapter 3). The CFG approach has been used extensively in studying the genetics of addiction and behavior disorders [122, 121]. The results our of genetic analysis were relatively unsurprising as all identified genes had been documented extensively as potentially associated with BD. The results of autoimmune disease analysis proved to be more fruitful as 16 autoimmune diseases were identified as having a genetic relationship to BD, 15 of which had documented associations with BD or case reports with patients presenting with both diseases simultaneously. These results are exciting for two reasons: 1) they indicate that the CFG approach can be used for autoimmune diseases and 2) they potentially indicate the presence of shared genetic pathways of the diseases. This presents an exciting area for future research as by unlocking potential similarities in the genetic pathways of autoimmune diseases, we can more effectively understand the patheogenesis of these diseases.

Finally, we took the top 5 genetically related autoimmune disease and analyzed their models to discover if they can be used for BD (Chapter 4). The most promising candidates were found in the HLA-A29 model, DC-EAU model, and P13K KO mouse model. All three of these models have previously not been used in BD research and represent exciting potential models that are able to present with a phenotypic symptom of BD in

conjunction with a genetic component. While these models are far from perfect representations of BD, they present potential opportunities for application in BD research.

FOOTNOTE

Portions of the data presented in this thesis were also submitted for publication. See references:

Shenoi, Samuel J, and Erich J Baker. "Using Hierarchical Similarity To Examine The Genetics of Behçet's Disease." *BioRxiv*, 2021. https://doi.org/10.1101/2021.04.06.438717.

Shenoi, Samuel J, and Erich J Baker. (2021). "*A review of Behçet's disease phenotype, genotype and animal models"* [Manuscript submitted for publication]. Department of Computer Science, Baylor University.

Shenoi, Samuel J, and Erich J Baker. (2021). "*Identification of An Improved Behçet's Disease Animal Model"* [Manuscript in preparation]. Department of Computer Science, Baylor University.

FOOTNOTE

Portions of the data presented in this thesis were also submitted for publication. See references:

Shenoi, Samuel J, and Erich J Baker. "Using Hierarchical Similarity To Examine The Genetics of Behçet's Disease." *BioRxiv*, 2021. https://doi.org/10.1101/2021.04.06.438717.

Shenoi, Samuel J, and Erich J Baker. (2021). "*A review of Behçet's disease phenotype, genotype and animal models"* [Manuscript submitted for publication]. Department of Computer Science, Baylor University.

Shenoi, Samuel J, and Erich J Baker. (2021). "*Identification of An Improved Behçet's Disease Animal Model"* [Manuscript in preparation]. Department of Computer Science, Baylor University.

APPENDICES

APPENDIX A

Uveitis Animal Models

Name	Species	Strain	Inoculation Technique	BD Symptoms
Transgenic HLA-A29 [349]	Mouse	Transgenic HLA-A29	Transgenic	retinopathy, posterior uveitis, retinal vasculitis, perivasculitis
Induced Model[347, 360]	Mouse, Rat	B10.R111, B10.A, C5BL/6, Lewis	ocular antigens in CFA or adoptive transfer of immune cells	posterior uveitis
Humanized Tg Models Mouse [347, 348]	Mouse	HLA-DR3	S-Ag in CFA posterior	chorditis, photoreceptor cell damage
RPE65 Rat [350, 347]	Rat	Lewis, Brown Norway (BN), Fischer (F344), and SHR rats	RPE65 in CFA and Bordetella pertussis injection	posterior uveitis, retinitis, engorgement of iridic blood vessels
DC-EAU Mouse [347, 351]	Mouse	B10.R111	infusion of antigen pulsed syngenic Dendritic Cells	posterior uveitis, engorged blood vessels

EMIU, EAAU [347, 352, 353]	Rat	pertussis toxin coadjuvant, 1 g of melanin-protein	TRP 1 and 2	anterior uveitis,neutrophil infiltrates in anterior chamber, retinal vasculitis
EAU bovine extract [353, 354, 355]	Rat, Monkey, Guinea Pig, Rabbit	various	immunized with a soluble bovine retinal extract	Uvea lesions, neutrophil lesions
EIU [347, 353, 356]	Mouse, Rat, Rabbit		Bacterial LPS	bilateral uveitis
AIRE knockout mice [347, 357]	Mouse	AIRE deficient mice	NA	Lymphocyte infiltrates in retina, salivary glands, and ovaries
AIRE lacking lymphocyte infected mice [357]	Mouse	alymphoid $Rag^{o/o}$	transference of lymphocytes from spleen and lymph nodes of AIRE deficient mice	Lymphocyte infiltrates in retina, salivary glands, ovaries,and stomach

Double Tg mice [347, 358]	Mouse	R161H mice	Double Tg mice expressing a neo self-antigen in the retina and an antigen specific TCR	uveitis
Experimental animal CMV retinitis model [347]	Mouse	C5BL/6	MCMV inoculation in MAIDS affected mice	retinitis
Experimental animal ocular tuberculosis model [347, 359]	Guinea pigs	Hartley strain	Mycobacterium tuberculosis inoculation	Uvea lesions

Table A.1: Uveitis Animal Models

APPENDIX B

Myositis Animal Models

Name	Species	Strain	Inoculation Technique	BD Symptoms
Diet-triggered rabbit myositis [363, 364]	Rabbit		Diet with 2% cholesterol	None
Transgenic Mice I [363, 365]	Mouse	C57BL/6 & SJL/J mice	transgenic MCK-APP/PS1 mice	None
Transgenic Mice II [363, 366]	Mouse	C57BL/6 & SJL/J mice	transgenic MLC-APP mice	None
Transgenic Mice III [363]	Mouse	Unknown	transgenic HSA-LTab mice	None
Transgenic Mice IV [363, 367]	Mouse	C57BL/6 (H+T+ mice)	MHC-I	None
Transgenic Mice V [363, 368]	Mouse	C57BL/6 (Syt VII /)	Synaptotagmin VII	Skin Collagen Increase
Transgenic Mice VI [369]	Mouse	BALB/c	K14-Angptl2	Eye and skin rashes

Transgenic Nematode [363]	C. Elegans	NA	transgenic Human A overexpressed in muscle	None
Canine Natural [363]	Dogs	NA	Natural	None
Infectious Mouse I [363, 370]	Mouse	BALB/c, Swiss C3H, COH mice	Coxsackievirus B1 infection	None
Infectious Mouse II [363, 371]	Mouse	CD-1, C57BL/6 mice	Ross River virus, Chikungunya virus (CHIKV SL15649) infection	Inflammatory cell arthritis and multifocal vasculitis
Infectious Mouse III [363]	Mouse	CBA/J mice	Trypanosoma cruzi infection	None
Infectious Immunological [363]	Mouse, Rat, Guinea Pig, Hamster	SJL/J mice, Lewis rats, Syrian Hamsters	Leishmania infantum Myosin B fraction infection	None
Immunological I [363, 372]	Rat, Mouse	Lewis rats, SJL/J, BALB/C, C57BL/6 mice	purified myosin	None

Immunological II [363, 373]	Rat, Mouse	Lewis rats, C57BL/6 mice	C-Protein fragment 2	None
Immunological III [363]	Mouse	C57BL/6	Transfer of CIM splenocytes cocultured with naïve BMDCs preincubated with CP2 and IL-2, simultaneous s.c. CFA injection.	None
Immunological IV [363, 374]	Rat	Wistar rats	Laminin	Inflammatory cell arthritis
Immunological V [363]	Mouse	C57BL/6, B6.G7, Nod.Idd3/5 mice	Amino terminal portion of HRS without CFA	None

Table B.1: Myositis Animal Models

APPENDIX C

Sarcoidosis Animal Models

Name	Species	Strain	Inoculation Technique	BD Symptoms
Kveim extract mice [401, 402, 403, 404, 405]	Mouse	various strains	footpad inoculation of Kveim sarcoidosis tissue extracts	None
CFA into Wistar rats [401, 406]	Rat	Wistar rats	injection of CFA	None
P. acnes/C.acnes in C57BL/6 mice [407, 401, 408, 409, 410, 411, 412]	Mouse	C57BL/6 mice	Propionibacterium acnes in C57BL/6 mice	None
M. tuberculosis Mouse [407, 413]	Mouse, Rat	C57BL/6, Lewis	M.tuberculosis infection	None
ApoE-deficient [408, 414]	Mouse		ApoE-deficient and high fat cholate diet	skin lesions and extra collagen in GI tract

Tsc2 [408, 415]	Mouse	myeloid lineage	tuberous sclerosis complex 2 (Tsc2) deletion	Skin lesions
Horse Sarcoidosis [416, 417, 418]	Horse		Natural	Skin lesions and GI tract involvement
Mycobacterial superoxide dismutase A [407, 419]	Mouse	C57BL/6 mice	Subcutaneous injections of IFA with incorporated mycobacterial sodA, followed by a challenge of beads-coupled sodA	None
H. pylori urease [407, 420]	Mouse	SD rats	preimmunized with intratracheal urease, and then challenged with H. pylori protein-bound beads intravenously 4 weeks later	None
Multiwall Carbon Nanotubes [407, 421]	Mouse	C57BL/6 mice	Oropharyngeal administration of multiwall carbon nanotubes	None

Table C.1: Sarcoidosis Animal Models

APPENDIX D

Sjogrens Animal Models

Name	Species	Strain	Inoculation Technique	BD Symptoms
Mouse Spontaneous I [376, 377]	Mouse	MRL/lpr, NZB/NZW F1	Spontaneous development of symptoms	Conjunctivitis, LG and SG lesions, increased neutrophil count, vasculitic infiltrates
Mouse Spontaneous II [376, 378, 379]	Mouse	Non-obese diabetic (NOD)	Spontaneous development of symptoms	Lesions of Salivary Glands and Testis, perivasciular infiltrates
Mouse Spontaneous III [376, 380, 378, 381, 382]	Mouse	C57BL/6.NOD-Aec1Aec2	Spontaneous development of symptoms	salivary glands and lacrimal gland lesions
Mouse Spontaneous IV [383, 376]	Mouse	NFS/sld, IQI/Jic	Spontaneous development of symptoms	salivary glands and lacrimal gland lesions
Mouse Spontaneous V [376, 384]	Mouse	Aly/aly	Spontaneous development of symptoms	salivary glands and lacrimal gland lesions

Mouse Id3 KO [376, 385]	Mouse	NA	Id3 KO	salivary and lacrimal gland lymphocytic infiltration
Mouse Aromatase KO [376, 386]	mouse	NA	Aromatase KO	salivary and lacrimal gland lymphocytic infiltration
Autoimmune regulator (Aire) KO mice [376, 387]	Mouse		Aire KO	salivary and lacrimal gland lymphocytic infiltration along with severe dry eye, stomach lymphoid cell infiltration
Mouse T cell-specific PI3K KO [376, 388]	Mouse	Lck-Cre crossed with r1f/r2n transgentic mice	T cell-specific PI3K KO	salivary and lacrimal gland lymphocytic infiltration, intestinal infiltration, eye lesions
IB$\alpha^{M/M}$ Model [376, 389]	Mouse	NA	IB$\alpha^{M/M}$ Knock In	salivary and lacrimal gland lymphocytic infiltration, perivascular infiltration
OPN transgenic [376, 390]	Mouse	Opn Tg mice backcrossed with C57BL/6	Osteopontin transgenic mice	salivary and lacrimal gland lymphocytic infiltration
Overexpression of RbAp48 in salivary gland [376, 391]	Mice	NA	RbAp48-Tg	salivary and lacrimal gland lymphocytic infiltration

Overexpression of TNF- in salivary gland [376, 392]	Mice	NA	AQP5-Cre transgenic	salivary gland lymphocytic infiltration
Ro immunisation (hRo60-480-494/274-290/273-289) [376, 393]	Mice	BALB/c	immunized with hRo60-480-494 or 274-290 or 273-289	salivary gland lymphocytic infiltration
Ro immunisation (mRo60-316-335)[376, 394]	Mice	C3H/HeJ, DBA/1J, C57BL/6J, and C3H/HeN	Murine $Ro60_3 16 - 335 immunization$	salivary and lacrimal gland lymphocytic infiltration
Ro immunisation (mRo52) [376]	Mice	New Zealand Mixed (NZM)	MBP-mRo52 immunization	salivary gland lymphocytic infiltration
M3R immunisation [376, 395]	Mice	C57BL/6j (B6) mice ($M3R^{+/+}$), $M3R^{-/-}$	immunized with a mixture of free-form extracellular peptides (each 20 g) in IFA with Mycobacterium tuberculosis and Pertussis toxin	salivary and lacrimal gland lymphocytic infiltration
IL-10 -/- mice [376, 465]	Mice	C57BL/6 NOD IL-10/ RAG-2/	Transgenic	salivary gland involvement

CA2 immunisation [376, 396]	Mice	PL/J (H-2u)	intradermal immunization with human carbonic anhydrase II (CAII) and adjuvant containing monophosphoryl lipid A and trehalose diorynomycolate	salivary and lacrimal gland lymphocytic infiltration
salivary gland infection of adenovirus 5 [376, 397]	Mice	C57BL/6	delivery of a replication-deficient adenovirus-5 through retrograde excretory duct cannulation	salivary gland lymphocytic infiltration
salivary gland cannulation of Ad5-IL17 vectors [376, 398]	Mice	C57BL/6J	Adenovirus serotype 5 (Ad5) vectors expressing either IL-17A or LacZ infused via retrograde cannulation	salivary gland lymphocytic infiltration
Chimeric human–mouse model [376, 399]	Mice	NSG	injection of isolated human PBMCs	salivary and lacrimall gland lymphocytic infiltration

Table D.1: Sjögren's Syndrome Animal Models

APPENDIX E

Hemolytic Anemia Animal Models

Name	Species	Strain	Inoculation Technique	BD Symptoms
New Zealand Black Mice (NZB) [424, 466]	Mice	New Zealand Black Mice	Spontaneous	None
Playfair and Marshall-Clarke model [424, 467]	Mice		intraperitoneal injection of rat RBCs	None
HL mice [424]	Mice	NA	Heavy and light chain derived from NZB model	None
HOD model [424, 468]	Mice	NA	BC-restricted triple fusion protein consisting of hen egg lysozyme (HEL), ovalbumin (OVA), and human blood group molecule	None

ATRX KO mice I [425]	Mice	NA	Atrx knockout mouse using the Cre-loxP recombination system controlled by the Forkhead box G1 (Foxg1; ref. 29) and Nestin promoter	None
ARTX KO mice II [426]	Mice	NA	Atrx knockout animals using the Cre-loxP recombination system in which Atrx ablation is restricted to the peripheral retina	Decreased horizontal and amacrine cells in retina
Glucose-6-Phosphate Isomerase Deficient Mice [427]	Mice	NA	$Gpi - 1s^{b-m1Neu}$ $Gpi - 1s^{b-m2Neu}$ mutants backcrossed with C3H/EI strain	None
nb/nb mice [428]	Mice	WB/Re and C57BL/6J mice	Spontaneous	Neurological Symptoms: unbalanced gait and moderate tremor
$ANK - 1^{1674}$ mutation[429]	Mice	C57BL/6	$ANK - 1^{1674}$ mutants outcrossed to C57BL/6 mice	None

ENU induced mutation in ankyrin-1 gene [430]	Mice	C57Bl/6J (B6), C3H/HeJ (C3H), and SvImJ/129 (129J) strains	three intraperitoneal injections of 85 mg/kg ENU	None
sph^{3J} and sph^{4J} mutations [431]	Mice	NOD/LtJ strain	spontaneous	None
sph^{2Bc} mutations [432]	Mice	Mixed strain: BALB/cGa, 129/J, and CBA/J, to the ICR/M1 stock	spontaneous	None
wan model [433]	Mice	C3H/HeJ (C3H)	spontaneous	None
PK deficient [434]	Mice	CBA-Pk-1^{slc}/Pk-1^{slc}	NA	None
Plasmodium chabaudi infection [435]	Mice	AcB55 X A/J, AcB61 x A/J, and AcB55 x DBA/2	Plasmodium chabaudi infection	None

EPO mice [436]	Mice	C57BL/6	treatment with erythiopoietin (EPO)	None
Polg D257A [437]	Mice	Polg D257A	Germline missense mutation in DNA polymerase gamma	None
AHSP-deficient mice [438]	Mice	α-Hb–stabilizing protein deficient	NA	None
Transgenic β-globin I [439]	Mice	C57BL/6	β-globin transgene	none
β-globin gene deletion II [440]	Mice	C57BL/6J(strain B6)	Mating between chimeras and C57BL/6J mice	None
Tmprss6 and Hbb$^{th3/+}$knockout mice [441]	Mice	NA	Tmprss6 and Hbb$^{th3/+}$ knockout	None
ELKF $^{-/-}$ [442]	Mice	NA	ELKF$^{-/-}$	None
RBC13 and RBC14 [443]	Mice	C57BL/6 and Balb/c	Mutation	None
M-II deficient [444]	Mice	C57BL/6	M-II deficient	None
mDia2 knockout mouse[445]	Mice	NA	mDia2 knockout	None

Nan mice [446]	Mice	C3H, C57BL/6J (B6), and WB/ReJ (WB)	mutation	None
Protein 4.2–null mice [447]	Mice	C57BL/6J	protein 4.2–null mice	None
β-adducin knockout mouse [448]	Mice	NA	β-adducin knockout	None
Gadd34-null mice [449]	Mice	NA	Gadd34-null mice	None
Radiation-induced [450]	Mice	Unknown	radiation-induced strain SEC alpha-chain deficiencies	None
Neo mouse[451]	Mice	α-globin PRE	replaced the AP1/NFE2 sites with a neomycin resistance gene (neo) that is flanked by LoxP sites (floxed)	None
Targeted deletion of the β-globin locus [452]	Mice	NA	targeted deletion of the β-globin locus control region	None

Insertional disruption of the mouse major adult β-globin gene [453]	Mice	NA	insertion of selectable sequences into β-globin gene	None
[454]	Mice	NA	replaced two (cis) murine adult β-globin genes with a single copy of the human βIVS-2-654 gene	None
Deletion of 5'HS2 of the murine β-globin LCR[455]	Mice	NA	Deletion of 5'HS2 of the murine β-globin LCR	None
Humanized Mouse Model of Cooley's Anemia [469]	Mice	C57BL/6J	($\gamma\beta^0$) insert	None
Transfusion-dependent humanized mouse [456]	Mice	C57BL/6J	mouse α and β globin genes replaced with $\alpha 2\alpha 1$ and $\delta\beta 0$,	None

Table E.1: Hemolytic Anemia animal models

APPENDIX F

Animal Models ICBD Score

Disease	Model	Oral Aphthosis	Genital Aphthosis	Ocular Lesions	Skin lesions	Neuro Involvement	GI Involvement	Vascular Involvement	Score	BD?
Uveitis/ Sjogrens	AIRE knockout								2	NO
Uveitis/ Sjogrens	AIRE lacking lympocyte infected mice								2	NO
Sarcoidosis	ApoE-deficient								1	NO
Sarcoidosis	Tsc2								1	NO
Sarcoidosis	Horse Sarcoidosis								1	NO
Myositis	Transgenic Mice V								0	NO
Myositis	Transgenic Mice VI								0	NO
Myositis	Infectious Mouse II								1	NO

Myositis	Immunological IV							0	NO
Uveitis	Transgenic HLA-A29							3	NO
Uveitis	Induced							2	NO
Uveitis	Humanized Tg							2	NO
Uveitis	RPE65 Rat							3	NO
Uveitis	DC-EAU							3	NO
Uveitis	”EMIU,EAAU”							3	NO
Uveitis	EAU bovine							2	NO
Uveitis	EIU							2	NO
Uveitis	Double Tg							2	NO
Uveitis	Experimental CMV retinitis							2	NO
Uveitis	Expermimental animal ocular tuberculosis							2	NO
Hemolytic Anemia	ARTX KO mice II							2	NO
Hemolytic Anemia	nb/nb							1	NO

Sjogrens	Mouse Spon-taneous I								3	NO
Sjogrens	Mouse Spon-taneous II								3	NO
Sjogrens	Mouse Spon-taneous III								0	NO
Sjogrens	Mouse Spon-taneous IV								0	NO
Sjogrens	Mouse Spon-taneous V								0	NO
Sjogrens	Mouse id3 KO								0	NO
Sjogrens	Mouse Aromatase KO								0	NO
Sjogrens	Mouse T cell specific P13K KO								2	NO

Sjogrens	IBM/M model								1	NO
Sjogrens	OPN trans-genic								0	NO
Sjogrens	Overexpression of RbAp48 in salivary gland								0	NO
Sjogrens	Overexpression of TNF-a in salivary gland								0	NO
Sjogrens	RO im-munosation								0	NO
Sjogrens	RO im-munosation II								0	NO
Sjogrens	RO im-munosation III								0	NO
Sjogrens	M3R im-munisation								0	NO
Sjogrens	Salivary Gland protein im-munisation								0	NO
Sjogrens	CA2 immu-nisation								0	NO

Sjogrens	salivary gland infection of adrenovirus								0	NO
Sjogrens	salivary gland cannulation of Ad5-IL17 vectors								0	NO
Sjogrens	Chimeric human-mouse model								0	NO

Table F.1: Resulting ICBD scores for each of the animal models identified with phenotypic presentation of symptoms in at least one BD associated area. None of the models scored above a 4, the minimum score required for a diagnosis, on the criteria.

Bibliography

[1] Davatchi, F. Behcet's disease. *International Journal of Rheumatic Diseases* **17**, 355–357 (2014). URL http://doi.wiley.com/10.1111/1756-185X.12378.

[2] Oğuz, A. K. *et al.* Behçet's: A Disease or a Syndrome? Answer from an Expression Profiling Study. *PloS One* **11**, e0149052 (2016).

[3] Bulur, I. & Onder, M. Behçet disease: New aspects. *Clinics in Dermatology* **35**, 421–434 (2017).

[4] Yildirim, O. Animal Models in Behçet's Disease. *Pathology Research International* **2012**, 273701 (2012).

[5] Baharav, E., Weinberger, A. & Mor, F. Experimental models of Behçet's disease. *Drug Discovery Today: Disease Models* **3**, 11–14 (2006). URL https://linkinghub.elsevier.com/retrieve/pii/S1740675706000090.

[6] Charles, J., Castellino, F. J. & Ploplis, V. A. Past and Present Behçet's Disease Animal Models. *Current Drug Targets* **21**, 1652–1663 (2020). URL https://www.eurekaselect.com/183880/article.

[7] Baker, E. J., Jay, J. J., Bubier, J. A., Langston, M. A. & Chesler, E. J. GeneWeaver: a web-based system for integrative functional genomics. *Nucleic Acids Research* **40**, D1067–1076 (2012).

[8] Baker, E., Bubier, J. A., Reynolds, T., Langston, M. A. & Chesler, E. J. GeneWeaver: data driven alignment of cross-species genomics in biology and disease. *Nucleic Acids Research* **44**, D555–D559 (2016). URL https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkv1329.

[9] Behçet, H. & Matteson, E. L. On relapsing, aphthous ulcers of the mouth, eye and genitalia caused by a virus. 1937. *Clinical and Experimental Rheumatology* **28**, S2–5 (2010).

[10] Kirino, Y. & Nakajima, H. Clinical and Genetic Aspects of Behçet's Disease in Japan. *Internal Medicine* **58**, 1199–1207 (2019). URL https://www.jstage.jst.go.jp/article/internalmedicine/58/9/58_2035-18/_article.

[11] Saccucci, M. *et al.* Autoimmune Diseases and Their Manifestations on Oral Cavity: Diagnosis and Clinical Management. *Journal of Immunology Research* **2018**, 6061825 (2018).

[12] Salmaninejad, A. *et al.* Genetics and immunodysfunction underlying Behçet's disease and immunomodulant treatment approaches. *Journal of Immunotoxicology* **14**, 137–151 (2017).

[13] Kalra, S. *et al.* Diagnosis and management of Neuro-Behçet's disease: international consensus recommendations. *Journal of Neurology* **261**, 1662–1676 (2014).

[14] Wu, Q.-J. Adamantiades-Behcet's disease-complicated gastroenteropathy. *World Journal of Gastroenterology* **18**, 609 (2012). URL http://www.wjgnet.com/1007-9327/full/v18/i7/609.htm.

[15] Tong, B., Liu, X., Xiao, J. & Su, G. Immunopathogenesis of Behcet's Disease. *Frontiers in Immunology* **10**, 665 (2019). URL https://www.frontiersin.org/article/10.3389/fimmu.2019.00665/full.

[16] International Team for the Revision of the International Criteria for Behçet's Disease (ITR-ICBD) *et al.* The International Criteria for Behçet's Disease (ICBD): a collaborative study of 27 countries on the sensitivity and specificity of the new criteria. *Journal of the European Academy of Dermatology and Venereology* **28**, 338–347 (2014). URL http://doi.wiley.com/10.1111/jdv.12107.

[17] Davatchi, F. Diagnosis/Classification Criteria for Behcet's Disease. *Pathology Research International* **2012**, 1–5 (2012). URL https://www.hindawi.com/archive/2012/607921/.

[18] Mizuki, N., Inoko, H. & Ohno, S. Molecular genetics (HLA) of Behçet's disease. *Yonsei Medical Journal* **38**, 333–349 (1997).

[19] HLA-B gene: MedlinePlus Genetics (2020). URL https://medlineplus.gov/genetics/gene/hla-b/.

[20] Janeway Jr, C. A., Travers, P., Walport, M. & Shlomchik, M. J. The major histocompatibility complex and its functions. In *Immunobiology: The Immune System in Health and Disease. 5th edition* (Garland Science, 2001).

[21] Verity, D., Marr, J., Ohno, S., Wallace, G. & Stanford, M. Behçet's disease, the Silk Road and HLA-B51: historical and geographical perspectives: Behçet's disease, the Silk Road and HLA-B51. *Tissue Antigens* **54**, 213–220 (1999). URL http://doi.wiley.com/10.1034/j.1399-0039.1999.540301.x.

[22] de Menthon, M., Lavalley, M. P., Maldini, C., Guillevin, L. & Mahr, A. HLA-B51/B5 and the risk of Behçet's disease: a systematic review and meta-analysis of case-control genetic association studies. *Arthritis and Rheumatism* **61**, 1287–1296 (2009).

[23] Rajaei, E. *et al.* Dose HLA-B5, 7, 8, 27, and 51 antigens associate to behcet's disease? A study in southwestern Iran. *Current Rheumatology Reviews* (2019).

[24] Mizuki, Y. *et al.* The influence of HLA-B51 on clinical manifestations among Japanese patients with Behçet's disease: A nationwide survey. *Modern Rheumatology* 1–7 (2019).

[25] Salvarani, C. *et al.* Association of MICA alleles and HLA-B51 in Italian patients with Behçet's disease. *The Journal of Rheumatology* **28**, 1867–1870 (2001).

[26] Muruganandam, M. *et al.* Characteristics of Behcet's Disease in the American Southwest. *Seminars in Arthritis and Rheumatism* **49**, 296–302 (2019). URL https://linkinghub.elsevier.com/retrieve/pii/S0049017218307911.

[27] Arber, N., Klein, T., Meiner, Z., Pras, E. & Weinberger, A. Close association of HLA-B51 and B52 in Israeli patients with Behcet's syndrome. *Annals of the Rheumatic Diseases* **50**, 351–353 (1991). URL https://ard.bmj.com/lookup/doi/10.1136/ard.50.6.351.

[28] Kilmartin, D. J., Finch, A. & Acheson, R. W. Primary association of HLA-B51 with Behcet's disease in Ireland. *British Journal of Ophthalmology* **81**, 649–653 (1997). URL https://bjo.bmj.com/lookup/doi/10.1136/bjo.81.8.649.

[29] Bodis, G., Toth, V. & Schwarting, A. Role of Human Leukocyte Antigens (HLA) in Autoimmune Diseases. *Rheumatology and Therapy* **5**, 5–20 (2018).

[30] Maldini, C., LaValley, M. P., Cheminant, M., de Menthon, M. & Mahr, A. Relationships of HLA-B51 or B5 genotype with Behçet's disease clinical characteristics: systematic review and meta-analyses of observational studies. *Rheumatology* **51**, 887–900 (2012). URL https://academic.oup.com/rheumatology/article-lookup/doi/10.1093/rheumatology/ker428.

[31] Giza, M., Koftori, D., Chen, L. & Bowness, P. Is Behçet's disease a 'class 1-opathy'? The role of HLA-B*51 in the pathogenesis of Behçet's disease. *Clinical and Experimental Immunology* **191**, 11–18 (2018).

[32] Gul, A. Lack of association of HLA-B*51 with a severe disease course in Behcet's disease. *Rheumatology* **40**, 668–672 (2001). URL https://academic.oup.com/rheumatology/article-lookup/doi/10.1093/rheumatology/40.6.668.

[33] Deng, Y., Zhu, W. & Zhou, X. Immune Regulatory Genes Are Major Genetic Factors to Behcet Disease: Systematic Review. *The Open Rheumatology Journal* **12**, 70–85 (2018). URL https://openrheumatologyjournal.com/VOLUME/12/PAGE/70/.

[34] Kaya, T. Genetics of Behçet's Disease. *Pathology Research International* **2012**, 912589 (2012).

[35] Shaker, O. G. *et al.* Expression of TNF- , April and BCMA in Behcet's Disease. *Journal of Immunology Research* **2014**, 1–6 (2014). URL http://www.hindawi.com/journals/jir/2014/380405/.

[36] Piga, M. & Mathieu, A. Genetic susceptibility to Behcet's disease: role of genes belonging to the MHC region. *Rheumatology* **50**, 299–310 (2011). URL https://academic.oup.com/rheumatology/article-lookup/doi/10.1093/rheumatology/keq331.

[37] Ahmad, T. *et al.* Mapping the HLA association in Behçet's disease: A role for tumor necrosis factor polymorphisms? *Arthritis & Rheumatism* **48**, 807–813 (2003). URL http://doi.wiley.com/10.1002/art.10815.

[38] Kalliolias, G. D. & Ivashkiv, L. B. TNF biology, pathogenic mechanisms and emerging therapeutic strategies. *Nature Reviews. Rheumatology* **12**, 49–62 (2016).

[39] Puccetti, A. *et al.* Gene Expression Profiling in Behcet's Disease Indicates an Autoimmune Component in the Pathogenesis of the Disease and Opens New Avenues for Targeted Therapy. *Journal of Immunology Research* **2018**, 1–18 (2018). URL https://www.hindawi.com/journals/jir/2018/4246965/.

[40] Takeuchi, M. *et al.* Dense genotyping of immune-related loci implicates host responses to microbial exposure in Behçet's disease susceptibility. *Nature Genetics* **49**, 438–443 (2017).

[41] Braschi, B. *et al.* Genenames.org: the HGNC and VGNC resources in 2019. *Nucleic Acids Research* **47**, D786–D792 (2019).

[42] Radouane, A. *et al.* SNPs in the TNF- gene promoter associated with Behcet's disease in Moroccan patients. *Rheumatology* **51**, 1595–1599 (2012). URL https://academic.oup.com/rheumatology/article-lookup/doi/10.1093/rheumatology/kes141.

[43] Lee, E. B., Kim, J. Y., Lee, Y. J., Park, M. H. & Song, Y. W. TNF and TNF receptor polymorphisms in Korean Behcet’s disease patients. *Human Immunology* **64**, 614–620 (2003). URL https://linkinghub.elsevier.com/retrieve/pii/S0198885903000570.

[44] Touma, Z. *et al.* TNF Polymorphisms in Patients with Behçet Disease: A Meta-analysis. *Archives of Medical Research* **41**, 142–146 (2010). URL https://linkinghub.elsevier.com/retrieve/pii/S0188440910000093.

[45] Zhang, M. *et al.* Polymorphisms in the tumor necrosis factor gene and susceptibility to Behcet's disease: an updated meta-analysis. *Molecular Vision* **19**, 1913–1924 (2013).

[46] Mizuki, N. *et al.* Triplet repeat polymorphism in the transmembrane region of the MICA gene: A strong association of six GCT repetitions with Behçet disease. *Proceedings of the National Academy of Sciences* **94**, 1298–1303 (1997). URL https://www.pnas.org/content/94/4/1298.

[47] Wang, Q. & Zhou, X. Associations of MICA Polymorphisms with Inflammatory Rheumatic Diseases. *The Open Rheumatology Journal* **9**, 94–100 (2015).

[48] Yabuki, K. *et al.* Association of MICA gene and HLA-B*5101 with Behçet's disease in Greece. *Investigative Ophthalmology & Visual Science* **40**, 1921–1926 (1999).

[49] Park, S. H. *et al.* Association of MICA Polymorphism with HLA-B51 and Disease Severity in Korean Patients with Behcet's Disease. *Journal of Korean Medical Science* **17**, 366–370 (2002). URL https://doi.org/10.3346/jkms.2002.17.3.366.

[50] Eyerci, N., Balkan, E., Akdeniz, N. & Keleş, S. Association of MICA Alleles and Human Leukocyte Antigen B in Turkish Patients Diagnosed With Behçet's Disease. *Archives of Rheumatology* **33**, 352–357 (2018).

[51] Picco, P. *et al.* MICA gene polymorphisms in an Italian paediatric series of juvenile Behçet disease. *International Journal of Molecular Medicine* **10**, 575–578 (2002).

[52] Wei, F., Zhang, Y. & Li, W. A meta-analysis of the association between Behçet's disease and MICA-A6. *Biomedical Reports* **4**, 741–745 (2016). URL https://www.spandidos-publications.com/10.3892/br.2016.644.

[53] Zhang, J., Liao, D., Yang, L. & Hou, S. Association between Functional MICA-TM and Behcet's Disease: A Systematic Review and Meta-analysis. *Scientific Reports* **6**, 21033 (2016). URL http://www.nature.com/articles/srep21033.

[54] Leccese, P. & Alpsoy, E. Behçet's Disease: An Overview of Etiopathogenesis. *Frontiers in Immunology* **10**, 1067 (2019).

[55] Takeuchi, M. *et al.* A single endoplasmic reticulum aminopeptidase-1 protein allotype is a strong risk factor for Behçet's disease in HLA-B*51 carriers. *Annals of the Rheumatic Diseases* **75**, 2208–2211 (2016). URL http://ard.bmj.com/lookup/doi/10.1136/annrheumdis-2015-209059.

[56] Conde-Jaldón, M. *et al.* Epistatic Interaction of ERAP1 and HLA-B in Behçet Disease: A Replication Study in the Spanish Population. *PLoS ONE* **9**, e102100 (2014). URL https://dx.plos.org/10.1371/journal.pone.0102100.

[57] Sousa, I. *et al.* Brief Report: Association of CCR1, KLRC4, IL12A–AS1, STAT4, and ERAP1 With Behçet's Disease in Iranians. *Arthritis & Rheumatology* **67**, 2742–2748 (2015). URL https://onlinelibrary.wiley.com/doi/abs/10.1002/art.39240.

[58] Tran, T. M. & Colbert, R. A. Endoplasmic reticulum aminopeptidase 1 and rheumatic disease: functional variation. *Current Opinion in Rheumatology* **27**, 357–363 (2015). URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002281-201507000-00007.

[59] Guasp, P. *et al.* The Behçet's disease-associated variant of the aminopeptidase ERAP1 shapes a low-affinity HLA-B*51 peptidome by differential subpeptidome processing. *Journal of Biological Chemistry* **292**, 9680–9689 (2017). URL http://www.jbc.org/lookup/doi/10.1074/jbc.M117.789180.

[60] López de Castro, J. A. How ERAP1 and ERAP2 Shape the Peptidomes of Disease-Associated MHC-I Proteins. *Frontiers in Immunology* **9**, 2463 (2018). URL https://www.frontiersin.org/article/10.3389/fimmu.2018.02463/full.

[61] Takeuchi, M., Kastner, D. L. & Remmers, E. F. The immunogenetics of Behçet's disease: A comprehensive review. *Journal of Autoimmunity* **64**, 137–148 (2015).

[62] Hou, S. *et al.* Identification of a susceptibility locus in STAT4 for Behçet's disease in Han Chinese in a genome-wide association study. *Arthritis & Rheumatism* **64**, 4104–4113 (2012). URL https://onlinelibrary.wiley.com/doi/abs/10.1002/art.37708.

[63] Kim, E. S. *et al.* Interactions between IL17A, IL23R, and STAT4 polymorphisms confer susceptibility to intestinal Behcet's disease in Korean population. *Life Sciences* **90**, 740–746 (2012). URL https://linkinghub.elsevier.com/retrieve/pii/S0024320512001403.

[64] Justiz Vaillant, A. A. & Qurie, A. Interleukin. In *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2021). URL http://www.ncbi.nlm.nih.gov/books/NBK499840/.

[65] Kim, S. W. *et al.* Identification of genetic susceptibility loci for intestinal Behçet's disease. *Scientific Reports* **7**, 39850 (2017).

[66] Remmers, E. F. *et al.* Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R-IL12RB2 regions associated with Behçet's disease. *Nature Genetics* **42**, 698–702 (2010).

[67] Hu, J. *et al.* Interleukin-10 gene polymorphisms are associated with Behcet's disease but not with Vogt-Koyanagi-Harada syndrome in the Chinese Han population. *Molecular Vision* **21**, 589–603 (2015).

[68] Aridogan, B. C. *et al.* Serum Levels of IL-4, IL-10, IL-12, IL-13 and IFN-gamma in Behçet's disease. *The Journal of Dermatology* **30**, 602–607 (2003).

[69] Turan, B. *et al.* Systemic levels of the T cell regulatory cytokines IL-10 and IL-12 in Bechçet's disease; soluble TNFR-75 as a biological marker of disease activity. *The Journal of Rheumatology* **24**, 128–132 (1997).

[70] Couper, K. N., Blount, D. G. & Riley, E. M. IL-10: The Master Regulator of Immunity to Infection. *The Journal of Immunology* **180**, 5771–5777 (2008). URL http://www.jimmunol.org/lookup/doi/10.4049/jimmunol.180.9.5771.

[71] Nakano, H. *et al.* GWAS-identified CCR1 and IL10 loci contribute to M1 macrophage-predominant inflammation in Behçet's disease. *Arthritis Research & Therapy* **20**, 124 (2018). URL https://arthritis-research.biomedcentral.com/articles/10.1186/s13075-018-1613-0.

[72] Shahriyari, E. *et al.* Exploring the association of IL-10 polymorphisms in Behcet's disease: a systematic review and meta-analysis. *Journal of Inflammation* **16**, 26 (2019). URL https://journal-inflammation.biomedcentral.com/articles/10.1186/s12950-019-0230-2.

[73] Li, X.-F., Huang, Z.-Z., Hu, Y.-Y., Niu, Y.-M. & Cai, J.-W. Association between Interleukin-10 Gene Polymorphisms and Behcet's Disease Susceptibility: Evidence from a Meta-Analysis. *Journal of Immunology Research* **2020**, 1–17 (2020). URL https://www.hindawi.com/journals/jir/2020/9382609/.

[74] Suzuki, N., Nara, K. & Suzuki, T. Skewed Th1 responses caused by excessive expression of Txk, a member of the Tec family of tyrosine kinases, in patients with Behcet's disease. *Clinical Medicine & Research* **4**, 147–151 (2006).

[75] Imamura, Y. *et al.* Involvement of Th1 cells and heat shock protein 60 in the pathogenesis of intestinal Behcet's disease. *Clinical and Experimental Immunology* **139**, 371–378 (2005).

[76] Shimizu, J., Kaneko, F. & Suzuki, N. Skewed Helper T-Cell Responses to IL-12 Family Cytokines Produced by Antigen-Presenting Cells and the Genetic Background in Behcet's Disease. *Genetics Research International* **2013**, 1–11 (2013). URL http://www.hindawi.com/journals/gri/2013/363859/.

[77] Chi, W. *et al.* Upregulated IL-23 and IL-17 in Behçet Patients with Active Uveitis. *Investigative Opthalmology & Visual Science* **49**, 3058 (2008). URL http://iovs.arvojournals.org/article.aspx?doi=10.1167/iovs.07-1390.

[78] Jiang, Z. *et al.* IL-23R gene confers susceptibility to Behcet's disease in a Chinese Han population. *Annals of the Rheumatic Diseases* **69**, 1325–1328 (2010).

[79] Yalçin, B., Atakan, N. & Dogan, S. Association of interleukin-23 receptor gene polymorphism with Behçet disease. *Clinical and Experimental Dermatology* **39**, 881–887 (2014). URL http://doi.wiley.com/10.1111/ced.12400.

[80] Zenobia, C. & Hajishengallis, G. Basic biology and role of interleukin-17 in immunity and inflammation. *Periodontology 2000* **69**, 142–159 (2015). URL http://doi.wiley.com/10.1111/prd.12083.

[81] Lopalco, G. *et al.* Cytokine Signatures in Mucocutaneous and Ocular Behçet's Disease. *Frontiers in Immunology* **8** (2017). URL http://journal.frontiersin.org/article/10.3389/fimmu.2017.00200/full.

[82] Nanke, Y., Yago, T. & Kotake, S. The Role of Th17 Cells in the Pathogenesis of Behcet's Disease. *Journal of Clinical Medicine* **6**, 74 (2017). URL http://www.mdpi.com/2077-0383/6/7/74.

[83] Burillo-Sanz, S. *et al.* Mutational profile of rare variants in inflammasome-related genes in Behçet disease: A Next Generation Sequencing approach. *Scientific Reports* **7**, 8453 (2017). URL http://www.nature.com/articles/s41598-017-09164-7.

[84] Goldberg, I., Baharav, E., Weinberger, A. & Krause, I. Animal Models of Behçet Syndrome. In Yazici, Y., Hatemi, G., Seyahi, E. & Yazici, H. (eds.) *Behçet Syndrome*, 235–242 (Springer International Publishing, Cham, 2020). URL http://link.springer.com/10.1007/978-3-030-24131-5_17.

[85] Sohn, S., Lee, E. S., Bang, D. & Lee, S. Behçet's disease-like symptoms induced by the Herpes simplex virus in ICR mice. *European journal of dermatology: EJD* **8**, 21–23 (1998).

[86] Eglin, R. P., Lehner, T. & Subak-Sharpe, J. H. Detection of RNA complementary to herpes-simplex virus in mononuclear cells from patients with Behçet's syndrome and recurrent oral ulcers. *Lancet (London, England)* **2**, 1356–1361 (1982).

[87] Islam, S. M. S. & Sohn, S. HSV-Induced Systemic Inflammation as an Animal Model for Behçet's Disease and Therapeutic Applications. *Viruses* **10** (2018).

[88] Sohn, S., Lee, E.-S. & Bang, D. Learning from HSV-infected mice as a model of Behçet's disease. *Clinical and Experimental Rheumatology* **30**, S96–103 (2012).

[89] Takeno, M. *et al.* Excessive function of peripheral blood neutrophils from patients with behcet's disease and from hla-b51 transgenic mice. *Arthritis & Rheumatism* **38**, 426–433 (1995). URL http://doi.wiley.com/10.1002/art.1780380321.

[90] Karaki, S. *et al.* HLA-B51 transgenic mice as recipients for production of polymorphic HLA-A, B-specific antibodies. *Immunogenetics* **37**, 139–142 (1993).

[91] Sato, Y., Nagata, S. & Takiguchi, M. Effective elicitation of human effector CD8+ T Cells in HLA-B*51:01 transgenic humanized mice after infection with HIV-1. *PloS One* **7**, e42776 (2012).

[92] Direskeneli, H. Innate and Adaptive Responses to Heat Shock Proteins in Behcet's Disease. *Genetics Research International* **2013**, 1–6 (2013). URL http://www.hindawi.com/journals/gri/2013/249157/.

[93] Sahebari, M., Hashemzadeh, K., Mahmoudi, M., Saremi, Z. & Mirfeizi, Z. Diagnostic Yield of Heat Shock Protein 70 (HSP-70) and Anti-HSP-70 in Behcet-Induced Uveitis. *Scandinavian Journal of Immunology* **77**, 476–481 (2013). URL http://doi.wiley.com/10.1111/sji.12045.

[94] Hu, W. *et al.* Experimental mucosal induction of uveitis with the 60-kDa heat shock protein-derived peptide 336-351. *European Journal of Immunology* **28**, 2444–2455 (1998).

[95] Galeone, M., Colucci, R., D'Erme, A. M., Moretti, S. & Lotti, T. Potential Infectious Etiology of Behçet's Disease. *Pathology Research International* **2012**, 595380 (2012).

[96] Namba, K., Ueno, T. & Okita, M. Behçet's disease and streptococcal infection. *Japanese Journal of Ophthalmology* **30**, 385–401 (1986).

[97] Cho, S. *et al.* Streptococcus sanguinis and the sera of patients with Behçet's disease stimulate membrane expression of -enolase in human dermal microvascular endothelial cells. *Archives of Dermatological Research* **305**, 223–232 (2013).

[98] Kaneko, F., Oyama, N. & Nishibu, A. Streptococcal infection in the pathogenesis of Behçet's disease and clinical effects of minocycline on the disease symptoms. *Yonsei Medical Journal* **38**, 444–454 (1997).

[99] Yüceer, H. & Tüzün, E. Animal Models of Behçet's Disease. In Tüzün, E. & Kürtüncü, M. (eds.) *Neuro-Behçet's Disease: Pathogenesis, Clinical Aspects, Treatment*, 41–48 (Springer International Publishing, Cham, 2021). URL https://doi.org/10.1007/978-3-030-55273-2_4.

[100] Mor, F., Weinberger, A. & Cohen, I. R. Identification of alpha-tropomyosin as a target self-antigen in Behçet's syndrome. *European Journal of Immunology* **32**, 356–365 (2002).

[101] Mahesh, S. P. *et al.* Alpha tropomyosin as a self-antigen in patients with Behçet's disease. *Clinical and Experimental Immunology* **140**, 368–375 (2005).

[102] Baharav, E., Mor, F., Halpern, M., Quintana, F. & Weinberger, A. Tropomyosin-induced arthritis in rats. *Clinical and Experimental Rheumatology* **25**, S86–92 (2007).

[103] Erdağ, E. *et al.* Effects of in vivo and in vitro administration of neuro-Behcet's disease IgG. *Neurological Sciences* **38**, 833–843 (2017). URL `http://link.springer.com/10.1007/s10072-017-2856-2`.

[104] McGonigle, P. & Ruggeri, B. Animal models of human disease: Challenges in enabling translation. *Biochemical Pharmacology* **87**, 162–171 (2014). URL `https://linkinghub.elsevier.com/retrieve/pii/S0006295213004929`.

[105] Barré-Sinoussi, F. & Montagutelli, X. Animal models are essential to biological research: issues and perspectives. *Future science OA* **1**, FSO63 (2015).

[106] Krishnan, V. & Nestler, E. J. Animal models of depression: molecular perspectives. *Current Topics in Behavioral Neurosciences* **7**, 121–147 (2011).

[107] Wang, Q., Timberlake, M. A., Prall, K. & Dwivedi, Y. The recent progress in animal models of depression. *Progress in Neuro-Psychopharmacology & Biological Psychiatry* **77**, 99–109 (2017).

[108] Eric W. Weisstein. Bipartite Graph – from Wolfram MathWorld. URL `https://mathworld.wolfram.com/BipartiteGraph.html`.

[109] Zhang, Y. *et al.* On finding bicliques in bipartite graphs: a novel algorithm and its application to the integration of diverse biological data types. *BMC bioinformatics* **15**, 110 (2014). Publisher: Springer.

[110] Bostock, M., Ogievetsky, V. & Heer, J. D^3 Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics* **17**, 2301–2309 (2011). URL `http://ieeexplore.ieee.org/document/6064996/`.

[111] Goyal, P. & Ferrara, E. Graph embedding techniques, applications, and performance: A survey. *Knowledge-Based Systems* **151**, 78–94 (2018). Publisher: Elsevier.

[112] Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* (2016).

[113] Kipf, T. N. & Welling, M. Variational Graph Auto-Encoders. *NIPS Workshop on Bayesian Deep Learning* (2016).

[114] Salha, G., Hennequin, R. & Vazirgiannis, M. Simple and effective graph autoencoders with one-hop linear models. *arXiv preprint arXiv:2001.07614* (2020).

[115] Fabian Pedregosa *et al.* Scikit-learn: Machine Learning in {P}ython}. *Journal of Machine Learning Research* **12**, 2825–2830 (2011). URL `http://jmlr.org/papers/v12/pedregosa11a.html`.

[116] Buitinck, L. *et al.* API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, 108–122 (2013).

[117] Hagberg, A. A., Schult, D. A. & Swart, P. J. Exploring Network Structure, Dynamics, and Function using NetworkX. In Varoquaux, G., Vaught, T. & Millman, J. (eds.) *Proceedings of the 7th Python in Science Conference*, 11 – 15 (Pasadena, CA USA, 2008).

[118] Paul Boutros. BoutrosLab.plotting.general: Functions to Create Publication-Quality Plots (2020). URL `https://CRAN.R-project.org/package=BoutrosLab.plotting.general`. R package version 6.0.1.

[119] R Core Team. R: A Language and Environment for Statistical Computing (2013). URL `http://www.R-project.org/`.

[120] Gibson, A. W., Edberg, J. C., Wu, J. & Kimberly, R. P. *The Role of IL-10 in Autoimmune Pathology* (Landes Bioscience, 2013). URL `https://www.ncbi.nlm.nih.gov/books/NBK6234/`. Publication Title: Madame Curie Bioscience Database [Internet].

[121] Spanagel, R. Convergent functional genomics in addiction research - a translational approach to study candidate genes and gene networks. *In Silico Pharmacology* **1**, 18 (2013).

[122] Niculescu, A. B. Convergent functional genomics of psychiatric disorders. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* **162**, 587–594 (2013). URL `https://onlinelibrary.wiley.com/doi/10.1002/ajmg.b.32163`.

[123] MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Research* **45**, D896–D901 (2017). URL `https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkw1133`.

[124] Lee, Y. J. *et al.* Genome-wide association study identifies GIMAP as a novel susceptibility locus for Behçet's disease. *Annals of the Rheumatic Diseases* **72**, 1510–1516 (2013). URL `http://ard.bmj.com/lookup/doi/10.1136/annrheumdis-2011-200288`.

[125] Fei, Y. *et al.* Identification of novel genetic susceptibility loci for Behçet's disease using a genome-wide association study. *Arthritis Research & Therapy* **11**, R66 (2009).

[126] Ortiz-Fernández, L. *et al.* Genetic Analysis with the Immunochip Platform in Behçet Disease. Identification of Residues Associated in the HLA Class I Region and New Susceptibility Loci. *PloS One* **11**, e0161305 (2016).

[127] Mizuki, N. *et al.* Genome-wide association studies identify IL23R-IL12RB2 and IL10 as Behçet's disease susceptibility loci. *Nature Genetics* **42**, 703–706 (2010).

[128] Kirino, Y. *et al.* Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLA-B*51 and ERAP1. *Nature Genetics* **45**, 202–207 (2013).

[129] Xavier, J. M. *et al.* *FUT2* : filling the gap between genes and environment in Behçet's disease? *Annals of the Rheumatic Diseases* **74**, 618–624 (2015). URL http://ard.bmj.com/lookup/doi/10.1136/annrheumdis-2013-204475.

[130] Kappen, J. H. *et al.* Genome-wide association study in an admixed case series reveals IL12A as a new candidate in Behçet disease. *PloS One* **10**, e0119085 (2015).

[131] Athey, B. D. *et al.* The NIH National Center for Integrative Biomedical Informatics (NCIBI). *Journal of the American Medical Informatics Association* **19**, 166–170 (2012). URL https://academic.oup.com/jamia/article-lookup/doi/10.1136/amiajnl-2011-000552.

[132] Amberger, J. S., Bocchini, C. A., Scott, A. F. & Hamosh, A. OMIM.org: leveraging knowledge across phenotype–gene relationships. *Nucleic Acids Research* **47**, D1038–D1043 (2019). URL https://academic.oup.com/nar/article/47/D1/D1038/5184722.

[133] Rappaport, N. *et al.* MalaCards: an integrated compendium for diseases and their annotation. *Database* **2013** (2013). URL https://academic.oup.com/database/article/doi/10.1093/database/bat018/331798.

[134] Köhler, S. *et al.* The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. *Nucleic Acids Research* **42**, D966–974 (2014).

[135] Birlea, S. A., Gowan, K., Fain, P. R. & Spritz, R. A. Genome-wide association study of generalized vitiligo in an isolated European founder population identifies SMOC2, in close proximity to IDDM8. *The Journal of Investigative Dermatology* **130**, 798–803 (2010).

[136] Cheong, K. A., Kim, N.-H., Noh, M. & Lee, A.-Y. Three new single nucleotide polymorphisms identified by a genome-wide association study in Korean patients with vitiligo. *Journal of Korean Medical Science* **28**, 775–779 (2013).

[137] Jin, Y. *et al.* Genome-wide analysis identifies a quantitative trait locus in the MHC class II region associated with generalized vitiligo age of onset. *The Journal of Investigative Dermatology* **131**, 1308–1312 (2011).

[138] Jin, Y. *et al.* Genome-wide association analyses identify 13 new susceptibility loci for generalized vitiligo. *Nature Genetics* **44**, 676–680 (2012).

[139] Jin, Y. *et al.* Variant of TYR and autoimmunity susceptibility loci in generalized vitiligo. *The New England Journal of Medicine* **362**, 1686–1697 (2010).

[140] Quan, C. *et al.* Genome-wide association study for vitiligo identifies susceptibility loci at 6q27 and the MHC. *Nature Genetics* **42**, 614–618 (2010).

[141] Tang, X.-F. *et al.* Association analyses identify three susceptibility Loci for vitiligo in the Chinese Han population. *The Journal of Investigative Dermatology* **133**, 403–410 (2013).

[142] Bracken, S., Byrne, G., Kelly, J., Jackson, J. & Feighery, C. Altered gene expression in highly purified enterocytes from patients with active coeliac disease. *BMC genomics* **9**, 377 (2008).

[143] Carr, E. J. *et al.* Confirmation of the genetic association of CTLA4 and PTPN22 with ANCA-associated vasculitis. *BMC medical genetics* **10**, 121 (2009).

[144] Curley, C. R., Monsuur, A. J., Wapenaar, M. C., Rioux, J. D. & Wijmenga, C. A functional candidate screen for coeliac disease genes. *European journal of human genetics: EJHG* **14**, 1215–1222 (2006).

[145] Diosdado, B. *et al.* A microarray screen for novel candidate genes in coeliac disease pathogenesis. *Gut* **53**, 944–951 (2004).

[146] Dubois, P. C. A. *et al.* Multiple common variants for celiac disease influencing immune gene expression. *Nature Genetics* **42**, 295–302 (2010).

[147] Festen, E. A. M. *et al.* A meta-analysis of genome-wide association scans identifies IL18RAP, PTPN2, TAGAP, and PUS10 as shared risk loci for Crohn's disease and celiac disease. *PLoS genetics* **7**, e1001283 (2011).

[148] Garner, C. *et al.* Genome-wide association study of celiac disease in North America confirms FRMD4B as new celiac locus. *PloS One* **9**, e101428 (2014).

[149] Hunt, K. A. *et al.* Newly identified genetic risk variants for celiac disease related to the immune response. *Nature Genetics* **40**, 395–402 (2008).

[150] Li, Y. R. *et al.* Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases. *Nature Medicine* **21**, 1018–1027 (2015).

[151] Östensson, M. *et al.* A possible mechanism behind autoimmune disorders discovered by genome-wide linkage and association analysis in celiac disease. *PloS One* **8**, e70174 (2013).

[152] Tello-Ruiz, M. K. *et al.* Haplotype-based association analysis of 56 functional candidate genes in the IBD6 locus on chromosome 19. *European journal of human genetics: EJHG* **14**, 780–790 (2006).

[153] van der Pouw Kraan, T. C. T. M., Zwiers, A., Mulder, C. J., Kraal, G. & Bouma, G. Acute experimental colitis and human chronic inflammatory diseases share expression of inflammation-related genes with conserved Ets2 binding sites. *Inflammatory Bowel Diseases* **15**, 224–235 (2009).

[154] van Heel, D. A. *et al.* A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. *Nature Genetics* **39**, 827–829 (2007).

[155] Zhernakova, A. *et al.* Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. *PLoS genetics* **7**, e1002004 (2011).

[156] Alonso, A. *et al.* Identification of risk loci for Crohn's disease phenotypes using a genome-wide association study. *Gastroenterology* **148**, 794–805 (2015).

[157] Barrett, J. C. *et al.* Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nature Genetics* **40**, 955–962 (2008).

[158] Dubinsky, M. C. *et al.* Multidimensional prognostic risk assessment identifies association between IL12B variation and surgery in Crohn's disease. *Inflammatory Bowel Diseases* **19**, 1662–1670 (2013).

[159] Ellinghaus, D. *et al.* Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci. *American Journal of Human Genetics* **90**, 636–647 (2012).

[160] Franke, A. *et al.* Genome-wide association analysis in sarcoidosis and Crohn's disease unravels a common susceptibility locus on 10p12.2. *Gastroenterology* **135**, 1207–1215 (2008).

[161] Franke, A. *et al.* Genome-wide association study for ulcerative colitis identifies risk loci at 7q22 and 22q13 (IL17REL). *Nature Genetics* **42**, 292–294 (2010).

[162] Huang, J., Ellinghaus, D., Franke, A., Howie, B. & Li, Y. 1000 Genomes-based imputation identifies novel and refined associations for the Wellcome Trust Case Control Consortium phase 1 Data. *European journal of human genetics: EJHG* **20**, 801–805 (2012).

[163] Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–124 (2012).

[164] Jung, E. S. *et al.* HLA-C*01 is a Risk Factor for Crohn's Disease. *Inflammatory Bowel Diseases* **22**, 796–806 (2016).

[165] Kenny, E. E. *et al.* A genome-wide scan of Ashkenazi Jewish Crohn's disease suggests novel susceptibility loci. *PLoS genetics* **8**, e1002559 (2012).

[166] Libioulle, C. *et al.* Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of PTGER4. *PLoS genetics* **3**, e58 (2007).

[167] McGovern, D. P. B. *et al.* Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn's disease. *Human Molecular Genetics* **19**, 3468–3476 (2010).

[168] Okada, Y. *et al.* HLA-Cw*1202-B*5201-DRB1*1502 haplotype increases risk for ulcerative colitis but reduces risk for Crohn's disease. *Gastroenterology* **141**, 864–871.e1–5 (2011).

[169] Parkes, M. *et al.* Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nature Genetics* **39**, 830–832 (2007).

[170] Raelson, J. V. *et al.* Genome-wide association study for Crohn's disease in the Quebec Founder Population identifies multiple validated disease loci. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 14747–14752 (2007).

[171] Rioux, J. D. *et al.* Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nature Genetics* **39**, 596–604 (2007).

[172] Yamazaki, K. *et al.* A genome-wide association study identifies 2 susceptibility Loci for Crohn's disease in a Japanese population. *Gastroenterology* **144**, 781–788 (2013).

[173] Yang, S.-K. *et al.* Genome-wide association study of Crohn's disease in Koreans revealed three new susceptibility loci and common attributes of genetic susceptibility across ethnic populations. *Gut* **63**, 80–87 (2014).

[174] Baurecht, H. *et al.* Genome-wide comparative analysis of atopic dermatitis and psoriasis gives insight into opposing genetic mechanisms. *American Journal of Human Genetics* **96**, 104–120 (2015).

[175] Capon, F. *et al.* Identification of ZNF313/RNF114 as a novel psoriasis susceptibility gene. *Human Molecular Genetics* **17**, 1938–1945 (2008).

[176] Ellinghaus, E. *et al.* Genome-wide association study identifies a psoriasis susceptibility locus at TRAF3IP2. *Nature Genetics* **42**, 991–995 (2010).

[177] Ellinghaus, E. *et al.* Genome-wide meta-analysis of psoriatic arthritis identifies susceptibility locus at REL. *The Journal of Investigative Dermatology* **132**, 1133–1140 (2012).

[178] Genetic Analysis of Psoriasis Consortium & the Wellcome Trust Case Control Consortium 2 *et al.* A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nature Genetics* **42**, 985–990 (2010).

[179] Haider, A. S. *et al.* Identification of cellular pathways of "type 1," Th17 T cells, and TNF- and inducible nitric oxide synthase-producing dendritic cells in autoimmune inflammation through pharmacogenomic study of cyclosporine A in psoriasis. *Journal of Immunology (Baltimore, Md.: 1950)* **180**, 1913–1920 (2008).

[180] Hüffmeier, U. *et al.* Common variants at TRAF3IP2 are associated with susceptibility to psoriatic arthritis and psoriasis. *Nature Genetics* **42**, 996–999 (2010).

[181] Li, Y. R. *et al.* Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases. *Nature Medicine* **21**, 1018–1027 (2015).

[182] Liu, Y. *et al.* A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. *PLoS genetics* **4**, e1000041 (2008).

[183] Nair, R. P. *et al.* Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. *Nature Genetics* **41**, 199–204 (2009).

[184] Stuart, P. E. *et al.* Genome-wide Association Analysis of Psoriatic Arthritis and Cutaneous Psoriasis Reveals Differences in Their Genetic Architecture. *American Journal of Human Genetics* **97**, 816–836 (2015).

[185] Stuart, P. E. *et al.* Genome-wide association analysis identifies three psoriasis susceptibility loci. *Nature Genetics* **42**, 1000–1004 (2010).

[186] Tsoi, L. C. *et al.* Enhanced meta-analysis and replication studies identify five new psoriasis susceptibility loci. *Nature Communications* **6**, 7001 (2015).

[187] Yin, X. *et al.* Genome-wide meta-analysis identifies multiple novel associations and ethnic heterogeneity of psoriasis susceptibility. *Nature Communications* **6**, 6916 (2015).

[188] Zhang, X.-J. *et al.* Psoriasis genome-wide association study identifies susceptibility variants within LCE gene cluster at 1q21. *Nature Genetics* **41**, 205–210 (2009).

[189] Awata, T. *et al.* A genome-wide association study for diabetic retinopathy in a Japanese population: potential association with a long intergenic non-coding RNA. *PloS One* **9**, e111715 (2014).

[190] Blackman, S. M. *et al.* Genetic modifiers of cystic fibrosis-related diabetes. *Diabetes* **62**, 3627–3635 (2013).

[191] Burdon, K. P. *et al.* Genome-wide association study for sight-threatening diabetic retinopathy reveals association with genetic variation near the GRB2 gene. *Diabetologia* **58**, 2288–2297 (2015).

[192] Germain, M. *et al.* SORBS1 gene, a new candidate for diabetic nephropathy: results from a multi-stage genome-wide association study in patients with type 1 diabetes. *Diabetologia* **58**, 543–548 (2015).

[193] Grassi, M. A. *et al.* Genome-wide meta-analysis for severe diabetic retinopathy. *Human Molecular Genetics* **20**, 2472–2481 (2011).

[194] Hanson, R. L. *et al.* A genome-wide association study in American Indians implicates DNER as a susceptibility locus for type 2 diabetes. *Diabetes* **63**, 369–376 (2014).

[195] Huang, Y.-C. *et al.* Genome-wide association study of diabetic retinopathy in a Taiwanese population. *Ophthalmology* **118**, 642–648 (2011).

[196] Imamura, M. *et al.* A single-nucleotide polymorphism in ANK1 is associated with susceptibility to type 2 diabetes in Japanese populations. *Human Molecular Genetics* **21**, 3042–3049 (2012).

[197] Iyengar, S. K. *et al.* Genome-Wide Association and Trans-ethnic Meta-Analysis for Advanced Diabetic Kidney Disease: Family Investigation of Nephropathy and Diabetes (FIND). *PLoS genetics* **11**, e1005352 (2015).

[198] Kwak, S. H. *et al.* A genome-wide association study of gestational diabetes mellitus in Korean women. *Diabetes* **61**, 531–541 (2012).

[199] Li, H. *et al.* A genome-wide association study identifies GRK5 and RASGRP1 as type 2 diabetes loci in Chinese Hans. *Diabetes* **62**, 291–298 (2013).

[200] Maeda, S. *et al.* Genetic variations associated with diabetic nephropathy and type II diabetes in a Japanese population. *Kidney International. Supplement* S43–48 (2007).

[201] McDonough, C. W. *et al.* A genome-wide association study for diabetic nephropathy genes in African Americans. *Kidney International* **79**, 563–572 (2011).

[202] Meng, W. *et al.* A Genome-wide Association Study Provides Evidence of Sex-specific Involvement of Chr1p35.1 (ZSCAN20-TLR12P) and Chr8p23.1 (HMGB1P46) With Diabetic Neuropathic Pain. *EBioMedicine* **2**, 1386–1393 (2015).

[203] Perry, J. R. B. *et al.* Stratifying type 2 diabetes cases by BMI identifies genetic risk variants in LAMA1 and enrichment for risk variants in lean compared to obese cases. *PLoS genetics* **8**, e1002741 (2012).

[204] Saxena, R. *et al.* Genome-wide association study identifies a novel locus contributing to type 2 diabetes susceptibility in Sikhs of Punjabi origin from India. *Diabetes* **62**, 1746–1755 (2013).

[205] SIGMA Type 2 Diabetes Consortium *et al.* Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature* **506**, 97–101 (2014).

[206] Tabassum, R. *et al.* Genome-wide association study for type 2 diabetes in Indians identifies a new susceptibility locus at 2q21. *Diabetes* **62**, 977–986 (2013).

[207] Takeuchi, F. *et al.* Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. *Diabetes* **58**, 1690–1699 (2009).

[208] Timpson, N. J. *et al.* Adiposity-related heterogeneity in patterns of type 2 diabetes susceptibility observed in genome-wide association data. *Diabetes* **58**, 505–510 (2009).

[209] Fischer, A. *et al.* A novel sarcoidosis risk locus for Europeans on chromosome 11q13.1. *American Journal of Respiratory and Critical Care Medicine* **186**, 877–885 (2012).

[210] Hofmann, S. *et al.* Genome-wide association analysis reveals 12q13.3-q14.1 as new risk locus for sarcoidosis. *The European Respiratory Journal* **41**, 888–900 (2013).

[211] Miller, F. W. *et al.* Genome-wide association study identifies HLA 8.1 ancestral haplotype alleles as major genetic risk factors for myositis phenotypes. *Genes and Immunity* **16**, 470–480 (2015).

[212] Allanore, Y. *et al.* Genome-wide scan identifies TNIP1, PSORS1C1, and RHOB as novel risk loci for systemic sclerosis. *PLoS genetics* **7**, e1002091 (2011).

[213] Gorlova, O. *et al.* Identification of novel genetic markers associated with clinical phenotypes of systemic sclerosis through a genome-wide association strategy. *PLoS genetics* **7**, e1002178 (2011).

[214] Martin, J.-E. *et al.* A systemic sclerosis and systemic lupus erythematosus pan-meta-GWAS reveals new shared susceptibility loci. *Human Molecular Genetics* **22**, 4021–4029 (2013).

[215] Radstake, T. R. D. J. *et al.* Genome-wide association study of systemic sclerosis identifies CD247 as a new susceptibility locus. *Nature Genetics* **42**, 426–429 (2010).

[216] Petukhova, L. *et al.* Genome-wide association study in alopecia areata implicates both innate and adaptive immunity. *Nature* **466**, 113–117 (2010).

[217] Schormair, B. *et al.* PTPRD (protein tyrosine phosphatase receptor type delta) is associated with restless legs syndrome. *Nature Genetics* **40**, 946–948 (2008).

[218] Stefansson, H. *et al.* A genetic risk factor for periodic limb movements in sleep. *The New England Journal of Medicine* **357**, 639–647 (2007).

[219] Winkelmann, J. *et al.* Genome-wide association study identifies novel restless legs syndrome susceptibility loci on 2p14 and 16q12.1. *PLoS genetics* **7**, e1002171 (2011).

[220] Winkelmann, J. *et al.* Genome-wide association study of restless legs syndrome identifies common variants in three genomic regions. *Nature Genetics* **39**, 1000–1006 (2007).

[221] Clancy, R. M. *et al.* Identification of candidate loci at 6p21 and 21q22 in a genome-wide association study of cardiac manifestations of neonatal lupus. *Arthritis and Rheumatism* **62**, 3415–3424 (2010).

[222] Ramos-Casals, M., Brito-Zeron, P., Siso-Almirall, A. & Bosch, X. Primary Sjogren syndrome. *BMJ* **344**, e3821–e3821 (2012). URL `http://www.bmj.com/cgi/doi/10.1136/bmj.e3821`.

[223] Song, I.-W. *et al.* Identification of susceptibility gene associated with female primary Sjögren's syndrome in Han Chinese by genome-wide association study. *Human Genetics* **135**, 1287–1294 (2016).

[224] Medici, M. *et al.* Identification of novel genetic Loci associated with thyroid peroxidase antibodies and clinical thyroid disease. *PLoS genetics* **10**, e1004123 (2014).

[225] Oryoji, D. *et al.* Identification of a Hashimoto thyroiditis susceptibility locus via a genome-wide comparison with Graves' disease. *The Journal of Clinical Endocrinology and Metabolism* **100**, E319–324 (2015).

[226] Duncan, E. L. *et al.* Genome-wide association study using extreme truncate selection identifies novel genes affecting bone mineral density and fracture risk. *PLoS genetics* **7**, e1001372 (2011).

[227] Australo-Anglo-American Spondyloarthritis Consortium (TASC) *et al.* Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nature Genetics* **42**, 123–127 (2010).

[228] Evans, D. M. *et al.* Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nature Genetics* **43**, 761–767 (2011).

[229] Lin, Z. *et al.* A genome-wide association study in Han Chinese identifies new susceptibility loci for ankylosing spondylitis. *Nature Genetics* **44**, 73–77 (2011).

[230] Deng, X. *et al.* Genome wide association study (GWAS) of Chagas cardiomyopathy in Trypanosoma cruzi seropositive subjects. *PloS One* **8**, e79629 (2013).

[231] Andlauer, T. F. M. *et al.* Novel multiple sclerosis susceptibility loci implicated in epigenetic regulation. *Science Advances* **2**, e1501678 (2016).

[232] Aulchenko, Y. S. *et al.* Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis. *Nature Genetics* **40**, 1402–1403 (2008).

[233] Australia and New Zealand Multiple Sclerosis Genetics Consortium (ANZgene). Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nature Genetics* **41**, 824–828 (2009).

[234] Baranzini, S. E. *et al.* Genetic variation influences glutamate concentrations in brains of patients with multiple sclerosis. *Brain: A Journal of Neurology* **133**, 2603–2611 (2010).

[235] Baranzini, S. E. *et al.* Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Human Molecular Genetics* **18**, 767–778 (2009).

[236] Brynedal, B. *et al.* MGAT5 alters the severity of multiple sclerosis. *Journal of Neuroimmunology* **220**, 120–124 (2010).

[237] Comabella, M. *et al.* Identification of a novel risk locus for multiple sclerosis at 13q31.3 by a pooled genome-wide scan of 500,000 single nucleotide polymorphisms. *PloS One* **3**, e3490 (2008).

[238] De Jager, P. L. *et al.* Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nature Genetics* **41**, 776–782 (2009).

[239] Goris, A. *et al.* Genetic variants are major determinants of CSF antibody levels in multiple sclerosis. *Brain: A Journal of Neurology* **138**, 632–643 (2015).

[240] Goris, A. *et al.* No evidence for shared genetic basis of common variants in multiple sclerosis and amyotrophic lateral sclerosis. *Human Molecular Genetics* **23**, 1916–1922 (2014).

[241] Gourraud, P.-A. *et al.* A genome-wide association study of brain lesion distribution in multiple sclerosis. *Brain: A Journal of Neurology* **136**, 1012–1024 (2013).

[242] International Multiple Sclerosis Genetics Consortium *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476**, 214–219 (2011).

[243] International Multiple Sclerosis Genetics Consortium. Genome-wide association study of severity in multiple sclerosis. *Genes and Immunity* **12**, 615–625 (2011).

[244] International Multiple Sclerosis Genetics Consortium *et al.* Risk alleles for multiple sclerosis identified by a genomewide study. *The New England Journal of Medicine* **357**, 851–862 (2007).

[245] Jakkula, E. *et al.* Genome-wide association study in a high-risk isolate for multiple sclerosis reveals associated variants in STAT3 gene. *American Journal of Human Genetics* **86**, 285–291 (2010).

[246] Leone, M. A. *et al.* Association of genetic markers with CSF oligoclonal bands in multiple sclerosis patients. *PloS One* **8**, e64408 (2013).

[247] Martinelli-Boneschi, F. *et al.* A genome-wide association study in progressive multiple sclerosis. *Multiple Sclerosis (Houndmills, Basingstoke, England)* **18**, 1384–1394 (2012).

[248] Matesanz, F. *et al.* Genome-wide association study of multiple sclerosis confirms a novel locus at 5p13.1. *PloS One* **7**, e36140 (2012).

[249] Mero, I.-L. *et al.* Oligoclonal band status in Scandinavian multiple sclerosis patients is associated with specific genetic risk alleles. *PloS One* **8**, e58352 (2013).

[250] Nischwitz, S. *et al.* Evidence for VAV2 and ZNF433 as susceptibility genes for multiple sclerosis. *Journal of Neuroimmunology* **227**, 162–166 (2010).

[251] Patsopoulos, N. A. *et al.* Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. *Annals of Neurology* **70**, 897–912 (2011).

[252] Sanna, S. *et al.* Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. *Nature Genetics* **42**, 495–497 (2010).

[253] Wang, J. H. *et al.* Modeling the cumulative genetic risk for multiple sclerosis from genome-wide association data. *Genome Medicine* **3**, 3 (2011).

[254] Zhou, Y. *et al.* Genetic loci for Epstein-Barr virus nuclear antigen-1 are associated with risk of multiple sclerosis. *Multiple Sclerosis (Houndmills, Basingstoke, England)* **22**, 1655–1664 (2016).

[255] Adachi, S. *et al.* Meta-analysis of genome-wide association scans for genetic susceptibility to endometriosis in Japanese population. *Journal of Human Genetics* **55**, 816–821 (2010).

[256] Albertsen, H. M., Chettier, R., Farrington, P. & Ward, K. Genome-wide association study link novel loci to endometriosis. *PloS One* **8**, e58257 (2013).

[257] Nyholt, D. R. *et al.* Genome-wide association meta-analysis identifies new endometriosis risk loci. *Nature Genetics* **44**, 1355–1359 (2012).

[258] Painter, J. N. *et al.* Genome-wide association study identifies a locus at 7p15.2 associated with endometriosis. *Nature Genetics* **43**, 51–54 (2011).

[259] Uno, S. *et al.* A genome-wide association study identifies genetic variants in the CDKN2BAS locus associated with endometriosis in Japanese. *Nature Genetics* **42**, 707–710 (2010).

[260] Wang, W. *et al.* Pooling-Based Genome-Wide Association Study Identifies Risk Loci in the Pathogenesis of Ovarian Endometrioma in Chinese Han Women. *Reproductive Sciences (Thousand Oaks, Calif.)* **24**, 400–406 (2017).

[261] Anderson, C. A. *et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nature Genetics* **43**, 246–252 (2011).

[262] Asano, K. *et al.* A genome-wide association study identifies three new susceptibility loci for ulcerative colitis in the Japanese population. *Nature Genetics* **41**, 1325–1329 (2009).

[263] Burczynski, M. E. *et al.* Molecular classification of Crohn's disease and ulcerative colitis patients using transcriptional profiles in peripheral blood mononuclear cells. *The Journal of molecular diagnostics: JMD* **8**, 51–61 (2006).

[264] Ellinghaus, D. *et al.* Genome-wide association analysis in primary sclerosing cholangitis and ulcerative colitis identifies risk loci at GPR35 and TCF4. *Hepatology (Baltimore, Md.)* **58**, 1074–1083 (2013).

[265] Flach, C.-F. *et al.* Detection of elafin as a candidate biomarker for ulcerative colitis by whole-genome microarray screening. *Inflammatory Bowel Diseases* **12**, 837–842 (2006).

[266] Franke, A. *et al.* Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. *Nature Genetics* **40**, 1319–1323 (2008).

[267] Haritunians, T. *et al.* Genetic predictors of medically refractory ulcerative colitis. *Inflammatory Bowel Diseases* **16**, 1830–1840 (2010).

[268] Julià, A. *et al.* A genome-wide association study identifies a novel locus at 6q22.1 associated with ulcerative colitis. *Human Molecular Genetics* **23**, 6927–6934 (2014).

[269] Juyal, G. *et al.* Genome-wide association scan in north Indians reveals three novel HLA-independent risk loci for ulcerative colitis. *Gut* **64**, 571–579 (2015).

[270] Liu, J. Z. *et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature Genetics* **47**, 979–986 (2015).

[271] McGovern, D. P. B. *et al.* Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nature Genetics* **42**, 332–337 (2010).

[272] Okahara, S. *et al.* Inflammatory gene signature in ulcerative colitis with cDNA macroarray analysis. *Alimentary Pharmacology & Therapeutics* **21**, 1091–1097 (2005).

[273] Silverberg, M. S. *et al.* Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nature Genetics* **41**, 216–220 (2009).

[274] UK IBD Genetics Consortium *et al.* Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. *Nature Genetics* **41**, 1330–1334 (2009).

[275] Waller, S. *et al.* Evidence for association of OCTN genes and IBD5 with ulcerative colitis. *Gut* **55**, 809–814 (2006).

[276] Watanabe, T. *et al.* Gene expression signature and the prediction of ulcerative colitis-associated colorectal cancer by DNA microarray. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* **13**, 415–420 (2007).

[277] Yang, S.-K. *et al.* Genome-wide association study of ulcerative colitis in Koreans suggests extensive overlapping of genetic susceptibility with Caucasians. *Inflammatory Bowel Diseases* **19**, 954–966 (2013).

[278] Zahn, A. *et al.* Aquaporin-8 expression is reduced in ileum and induced in colon of patients with ulcerative colitis. *World Journal of Gastroenterology* **13**, 1687–1695 (2007).

[279] Milton, J. N. *et al.* A genome-wide association study of total bilirubin and cholelithiasis risk in sickle cell anemia. *PloS One* **7**, e34741 (2012).

[280] Alarcón-Riquelme, M. E. *et al.* Genome-Wide Association Study in an Amerindian Ancestry Population Reveals Novel Systemic Lupus Erythematosus Risk Loci and the Role of European Admixture. *Arthritis & Rheumatology (Hoboken, N.J.)* **68**, 932–943 (2016).

[281] Armstrong, D. L. *et al.* GWAS identifies novel SLE susceptibility genes and explains the association of the HLA region. *Genes and Immunity* **15**, 347–354 (2014).

[282] Bentham, J. *et al.* Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nature Genetics* **47**, 1457–1464 (2015).

[283] Chung, S. A. *et al.* Differential genetic associations for systemic lupus erythematosus based on anti-dsDNA autoantibody production. *PLoS genetics* **7**, e1001323 (2011).

[284] Demirci, F. Y. *et al.* Identification of a New Susceptibility Locus for Systemic Lupus Erythematosus on Chromosome 12 in Individuals of European Ancestry. *Arthritis & Rheumatology (Hoboken, N.J.)* **68**, 174–183 (2016).

[285] Graham, R. R. *et al.* Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. *Nature Genetics* **40**, 1059–1061 (2008).

[286] Han, J.-W. *et al.* Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nature Genetics* **41**, 1234–1237 (2009).

[287] Hom, G. *et al.* Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. *The New England Journal of Medicine* **358**, 900–909 (2008).

[288] International Consortium for Systemic Lupus Erythematosus Genetics (SLEGEN) *et al.* Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. *Nature Genetics* **40**, 204–210 (2008).

[289] Kariuki, S. N. *et al.* Genetic analysis of the pathogenic molecular sub-phenotype interferon-alpha identifies multiple novel loci involved in systemic lupus erythematosus. *Genes and Immunity* **16**, 15–23 (2015).

[290] Kim, K. *et al.* Response to Intravenous Cyclophosphamide Treatment for Lupus Nephritis Associated with Polymorphisms in the FCGR2B-FCRLA Locus. *The Journal of Rheumatology* **43**, 1045–1049 (2016).

[291] Kozyrev, S. V. *et al.* Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. *Nature Genetics* **40**, 211–216 (2008).

[292] Lee, Y. H., Bae, S.-C., Choi, S. J., Ji, J. D. & Song, G. G. Genome-wide pathway analysis of genome-wide association studies on systemic lupus erythematosus and rheumatoid arthritis. *Molecular Biology Reports* **39**, 10627–10635 (2012).

[293] Lessard, C. J. *et al.* Identification of a Systemic Lupus Erythematosus Risk Locus Spanning ATG16L2, FCHSD2, and P2RY2 in Koreans. *Arthritis & Rheumatology (Hoboken, N.J.)* **68**, 1197–1209 (2016).

[294] Márquez, A. *et al.* A combined large-scale meta-analysis identifies COG6 as a novel shared risk locus for rheumatoid arthritis and systemic lupus erythematosus. *Annals of the Rheumatic Diseases* **76**, 286–294 (2017).

[295] Morris, D. L. *et al.* Genome-wide association meta-analysis in Chinese and European individuals identifies ten new loci associated with systemic lupus erythematosus. *Nature Genetics* **48**, 940–946 (2016).

[296] Okada, Y. *et al.* A genome-wide association study identified AFF1 as a susceptibility locus for systemic lupus eyrthematosus in Japanese. *PLoS genetics* **8**, e1002455 (2012).

[297] Yang, J. *et al.* ELF1 is associated with systemic lupus erythematosus in Asian populations. *Human Molecular Genetics* **20**, 601–607 (2011).

[298] Yang, W. *et al.* Genome-wide association study in Asian populations identifies variants in ETS1 and WDFY4 associated with systemic lupus erythematosus. *PLoS genetics* **6**, e1000841 (2010).

[299] Zhang, Y. *et al.* Genome-wide search followed by replication reveals genetic interaction of CD80 and ALOX5AP associated with systemic lupus erythematosus in Asian populations. *Annals of the Rheumatic Diseases* **75**, 891–898 (2016).

[300] Cordell, H. J. *et al.* International genome-wide meta-analysis identifies new primary biliary cirrhosis risk loci and targetable pathogenic pathways. *Nature Communications* **6**, 8019 (2015).

[301] Garcia-Barceló, M.-M. *et al.* Genome-wide association study identifies a susceptibility locus for biliary atresia on 10q24.2. *Human Molecular Genetics* **19**, 2917–2925 (2010).

[302] Hirschfield, G. M. *et al.* Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. *The New England Journal of Medicine* **360**, 2544–2555 (2009).

[303] Liu, X. *et al.* Genome-wide meta-analyses identify three loci associated with primary biliary cirrhosis. *Nature Genetics* **42**, 658–660 (2010).

[304] Mells, G. F. *et al.* Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. *Nature Genetics* **43**, 329–332 (2011).

[305] Nakamura, M. *et al.* Genome-wide association study identifies TNFSF15 and POU2AF1 as susceptibility loci for primary biliary cirrhosis in the Japanese population. *American Journal of Human Genetics* **91**, 721–728 (2012).

[306] Chen, P.-L. *et al.* Genetic determinants of antithyroid drug-induced agranulocytosis by human leukocyte antigen genotyping and genome-wide association study. *Nature Communications* **6**, 7633 (2015).

[307] Chu, X. *et al.* A genome-wide association study identifies two new risk loci for Graves' disease. *Nature Genetics* **43**, 897–901 (2011).

[308] Khong, J. J. *et al.* Association of Polymorphisms in MACRO Domain Containing 2 With Thyroid-Associated Orbitopathy. *Investigative Ophthalmology & Visual Science* **57**, 3129–3137 (2016).

[309] Nakabayashi, K. *et al.* Identification of independent risk loci for Graves' disease within the MHC in the Japanese population. *Journal of Human Genetics* **56**, 772–778 (2011).

[310] Zhao, S.-X. *et al.* Robust evidence for five new Graves' disease risk loci from a staged genome-wide association analysis. *Human Molecular Genetics* **22**, 3347–3362 (2013).

[311] Feehally, J. *et al.* HLA has strongest association with IgA nephropathy in genome-wide analysis. *Journal of the American Society of Nephrology: JASN* **21**, 1791–1797 (2010).

[312] Gharavi, A. G. *et al.* Genome-wide association study identifies susceptibility loci for IgA nephropathy. *Nature Genetics* **43**, 321–327 (2011).

[313] Kiryluk, K. *et al.* Discovery of new risk loci for IgA nephropathy implicates genes involved in immunity against intestinal pathogens. *Nature Genetics* **46**, 1187–1196 (2014).

[314] Li, M. *et al.* Identification of new susceptibility loci for IgA nephropathy in Han Chinese. *Nature Communications* **6**, 7270 (2015).

[315] Yang, C. *et al.* Genome-wide association study identifies TNFSF13 as a susceptibility gene for IgA in a South Chinese population in smokers. *Immunogenetics* **64**, 747–753 (2012).

[316] Yu, X.-Q. *et al.* A genome-wide association study in Han Chinese identifies multiple susceptibility loci for IgA nephropathy. *Nature Genetics* **44**, 178–182 (2011).

[317] Gabor Csardi & Tamas Nepusz. The igraph software package for complex network research. *InterJournal* **Complex Systems**, 1695 (2006). URL `http://igraph.org`.

[318] E. Paradis & K. Schliep. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in {R}. *Bioinformatics* **35**, 526–528 (2019).

[319] Salvatore, S. *et al.* Beware the Jaccard: the choice of **similarity measure** is important and non-trivial in genomic colocalisation analysis. *Briefings in Bioinformatics* **21**, 1523–1530 (2020). URL `https://academic.oup.com/bib/article/21/5/1523/5586919`.

[320] Nair, J. & Singh, T. Sjogren's syndrome: Review of the aetiology, Pathophysiology a Potential therapeutic interventions. *Journal of Clinical and Experimental Dentistry* 0–0 (2017). URL `http://www.medicinaoral.com/medoralfree01/aop/53605.pdf`.

[321] Günaydin, I. *et al.* The prevalence of Sjögren's syndrome in Behçet's syndrome. *The Journal of Rheumatology* **21**, 1662–1664 (1994).

[322] Ju, F.-H. *et al.* Behcet disease combined with Sjogren syndrome: A unique case report and literature review. *Medicine* **97**, e0138 (2018).

[323] Louro, M., Vaio, T., Crespo, J., Santos, R. & Carvalho, A. Pulmonary Sarcoidosis in Behçet's Disease Treated with Adalimumab. *European Journal of Case Reports in Internal Medicine* **4**, 000576 (2017).

[324] Berriche, O. *et al.* Behçet's disease and sarcoidosis: a rare association. *Research* **1** (2014). URL `http://www.labome.org/research/Beh-et-s-disease-and-sarcoidosis-a-rare-association.html`.

[325] Kokturk, A. Clinical and Pathological Manifestations with Differential Diagnosis in Behçet's Disease. *Pathology Research International* **2012**, 1–9 (2012). URL `https://www.hindawi.com/journals/pri/2012/690390/`.

[326] Hahn, H. J., Kwak, S. G., Kim, D.-K. & Kim, J.-Y. Association of Behçet disease with psoriasis and psoriatic arthritis. *Scientific Reports* **11**, 2531 (2021). URL `http://www.nature.com/articles/s41598-021-81972-4`.

[327] Nordstrom, E. & Fischer, M. The great masquerader: Behcet's disease. *Case Reports* **2014**, bcr2013202919–bcr2013202919 (2014). URL https://casereports.bmj.com/lookup/doi/10.1136/bcr-2013-202919.

[328] Watanabe, H. *et al.* A CASE OF BEHÇET'S DISEASE AND SYSTEMIC SCLEROSIS DEVELOPING AFTER AN EARTHQUAKE DISASTER. *FUKUSHIMA JOURNAL OF MEDICAL SCIENCE* **61**, 86–90 (2015). URL https://www.jstage.jst.go.jp/article/fms/61/1/61_2014-27/_article.

[329] Ichimura, Y. *et al.* Development of systemic sclerosis in patients with Behçet's disease: Remission of Behçet's disease in parallel with the progression of skin sclerosis. *The Journal of Dermatology* **41**, 1113–1114 (2014). URL http://doi.wiley.com/10.1111/1346-8138.12688.

[330] Ogose, T. *et al.* A case of recurrent myositis as the main manifestation of Behçet disease. *Pediatrics International* **52**, e101–e104 (2010). URL http://doi.wiley.com/10.1111/j.1442-200X.2010.03082.x.

[331] Dursun, D., Akova, Y. & Yücel, E. Myositis and scleritis associated with Behcet's disease: an atypical presentation. *Ocular Immunology and Inflammation* **12**, 329–332 (2004).

[332] Sarui, H. Necrotising myositis in Behcet's disease: characteristic features on magnetic resonance imaging and a review of the literature. *Annals of the Rheumatic Diseases* **61**, 751–752 (2002). URL https://ard.bmj.com/lookup/doi/10.1136/ard.61.8.751.

[333] Carlos Arteaga, R., Otto J Hernandez, F., Renato Puppi, M. & Olga Judith Hernández, F. Polymyositis in Adamantiades-Behçet's Disease. *International Journal of Neurology and Neurotherapy* **7** (2020). URL https://www.clinmedjournals.org/articles/ijnn/international-journal-of-neurology-and-neurotherapy-ijnn-7-094.php?jid=ijnn.

[334] Mignemi, G. *et al.* A Case Report of Nasal Endometriosis in a Patient Affected by Behcet's Disease. *Journal of Minimally Invasive Gynecology* **19**, 514–516 (2012). URL https://linkinghub.elsevier.com/retrieve/pii/S1553465012001082.

[335] Joshi, H., Shahriar, I., Sharma, P., Sagi, S. V. & Oyibo, S. O. A rare coexistence of Behcet's disease and Graves' thyrotoxicosis in a young man: a case report. *Oxford Medical Case Reports* **2020**, omz132 (2020). URL https://academic.oup.com/omcr/article/doi/10.1093/omcr/omz132/5669843.

[336] Ertaş, R., Özyurt, K., Avcı, A., Ketenci Ertas, S. & Atasoy, M. Case Report: Behçet's disease accompanied with vitiligo. *F1000Research* **6**, 310 (2017). URL https://f1000research.com/articles/6-310/v1.

[337] Guney, E. *et al.* Vitiligo in a Patient Treated with Interferon Alpha-2a for Behçet's Disease. *Case Reports in Medicine* **2012**, 1–3 (2012). URL http://www.hindawi.com/journals/crim/2012/387140/.

[338] Chyuma, Y. *et al.* [Hemolytic anemia complicated with Behçet's disease and myelodysplastic syndrome]. *[Rinsho Ketsueki] The Japanese Journal of Clinical Hematology* **33**, 333–337 (1992).

[339] Chung, S. Y., Oh, K. K. & Chang, H. S. Sonographic findings of tuberculous thyroiditis in a patient with Behçet's syndrome: Tuberculous Thyroiditis. *Journal of Clinical Ultrasound* **30**, 184–188 (2002). URL http://doi.wiley.com/10.1002/jcu.10041.

[340] Carvalho, F. P. *et al.* Subacute thyroiditis in association with psoriasis and behcet's disease. *Endocrine Abstracts* (2018). URL http://www.endocrine-abstracts.org/ea/0056/ea0056p635.htm.

[341] Iwadate, H. A case of primary biliary cirrhosis complicated by Behçet's disease and palmoplantar pustulosis. *World Journal of Gastroenterology* **12**, 2136 (2006). URL http://www.wjgnet.com/1007-9327/full/v12/i13/2136.htm.

[342] Shenoi, S. J. & Baker, E. Behcet's Disease Network Analysis: A Convergent Functional Genomics Approach. preprint, Bioinformatics (2021). URL http://biorxiv.org/lookup/doi/10.1101/2021.04.06.438717.

[343] Bult, C. J. *et al.* Mouse Genome Database (MGD) 2019. *Nucleic Acids Research* **47**, D801–D806 (2019). URL https://academic.oup.com/nar/article/47/D1/D801/5165331.

[344] Hsu, Y.-R. *et al.* Noninfectious uveitis in the Asia-Pacific region. *Eye (London, England)* **33**, 66–77 (2019).

[345] Rosenbaum, J. T. & Dick, A. D. The Eyes Have it: A Rheumatologist's View of Uveitis. *Arthritis & Rheumatology (Hoboken, N.J.)* **70**, 1533–1543 (2018).

[346] Forrester, J. V., Kuffova, L. & Dick, A. D. Autoimmunity, Autoinflammation, and Infection in Uveitis. *American Journal of Ophthalmology* **189**, 77–85 (2018). URL https://linkinghub.elsevier.com/retrieve/pii/S0002939418300941.

[347] Agrawal, R., Barathi, V., Iwata, D. & Bansal, S. Experimental autoimmune uveitis and other animal models of uveitis: An update. *Indian Journal of Ophthalmology* **63**, 211 (2015). URL http://www.ijo.in/text.asp?2015/63/3/211/156914.

[348] Pennesi, G. *et al.* A humanized model of experimental autoimmune uveitis in HLA class II transgenic mice. *The Journal of Clinical Investigation* **111**, 1171–1180 (2003).

[349] Szpak, Y. *et al.* Spontaneous retinopathy in HLA-A29 transgenic mice. *Proceedings of the National Academy of Sciences* **98**, 2572–2576 (2001). URL http://www.pnas.org/cgi/doi/10.1073/pnas.051595998.

[350] Ham, D.-I. *et al.* RPE65 is highly uveitogenic in rats. *Investigative Ophthalmology & Visual Science* **43**, 2258–2263 (2002).

[351] Tang, J. *et al.* Autoimmune uveitis elicited with antigen-pulsed dendritic cells has a distinct clinical signature and is driven by unique effector mechanisms: initial encounter with autoantigen defines disease phenotype. *Journal of Immunology (Baltimore, Md.: 1950)* **178**, 5578–5587 (2007).

[352] Broekhuyse, R., Kuhlmann, E. & Winkens, H. Experimental autoimmune anterior uveitis (EAAU). II. Dose-dependent induction and adoptive transfer using a melanin-bound antigen of the retinal pigment epithelium. *Experimental Eye Research* **55**, 401–411 (1992). URL `https://linkinghub.elsevier.com/retrieve/pii/0014483592901126`.

[353] Smith, J. R., Hart, P. H. & Williams, K. A. Basic pathogenic mechanisms operating in experimental models of acute anterior uveitis. *Immunology and Cell Biology* **76**, 497–512 (1998).

[354] Wacker, W. B. & Lipton, M. M. Experimental allergic uveitis: homologous retina as uveitogenic antigen. *Nature* **206**, 253–254 (1965).

[355] Wacker, W. B. Proctor Lecture. Experimental allergic uveitis. Investigations of retinal autoimmunity and the immunopathologic responses evoked. *Investigative Ophthalmology & Visual Science* **32**, 3119–3128 (1991).

[356] Rosenbaum, J. T., McDevitt, H. O., Guss, R. B. & Egbert, P. R. Endotoxin-induced uveitis in rats as a model for human disease. *Nature* **286**, 611–613 (1980).

[357] Anderson, M. S. *et al.* Projection of an immunological self shadow within the thymus by the aire protein. *Science (New York, N.Y.)* **298**, 1395–1401 (2002).

[358] Horai, R. *et al.* Breakdown of immune privilege and spontaneous autoimmunity in mice expressing a transgenic T cell receptor specific for a retinal autoantigen. *Journal of Autoimmunity* **44**, 21–33 (2013).

[359] Rao, N. A. *et al.* Experimental ocular tuberculosis in guinea pigs. *Archives of Ophthalmology (Chicago, Ill.: 1960)* **127**, 1162–1166 (2009).

[360] Caspi, R. R., Grubbs, B. G., Chan, C. C., Chader, G. J. & Wiggert, B. Genetic control of susceptibility to experimental autoimmune uveoretinitis in the mouse model. Concomitant regulation by MHC and non-MHC genes. *Journal of Immunology (Baltimore, Md.: 1950)* **148**, 2384–2389 (1992).

[361] Dix, R. D., Cray, C. & Cousins, S. W. Mice immunosuppressed by murine retrovirus infection (MAIDS) are susceptible to cytomegalovirus retinitis. *Current Eye Research* **13**, 587–595 (1994).

[362] Schmidt, J. Current Classification and Management of Inflammatory Myopathies. *Journal of Neuromuscular Diseases* **5**, 109–129 (2018). URL `https://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/JND-180308`.

[363] Afzali, A. M., Ruck, T., Wiendl, H. & Meuth, S. G. Animal models in idiopathic inflammatory myopathies: How to overcome a translational roadblock? *Autoimmunity Reviews* **16**, 478–494 (2017). URL `https://linkinghub.elsevier.com/retrieve/pii/S156899721730054X`.

[364] Sparks, D. *et al.* Induction of Alzheimer-like -Amyloid Immunoreactivity in the Brains of Rabbits with Dietary Cholesterol. *Experimental Neurology* **126**, 88–94 (1994). URL https://linkinghub.elsevier.com/retrieve/pii/S0014488684710442.

[365] Sugarman, M. C. *et al.* Inclusion body myositis-like phenotype induced by transgenic overexpression of beta APP in skeletal muscle. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 6334–6339 (2002).

[366] E-H. Moussa, C. *et al.* Transgenic expression of B-APP in fast-twitch skeletal muscle leads to calcium dyshomeostasis and IBM-like pathology. *The FASEB Journal* **20**, 2165–2167 (2006). URL https://onlinelibrary.wiley.com/doi/abs/10.1096/fj.06-5763fje.

[367] Nagaraju, K. *et al.* Conditional up-regulation of MHC class I in skeletal muscle leads to self-sustaining autoimmune myositis and myositis-specific autoantibodies. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 9209–9214 (2000).

[368] Chakrabarti, S. *et al.* Impaired membrane resealing and autoimmune myositis in synaptotagmin VII–deficient mice. *Journal of Cell Biology* **162**, 543–549 (2003). URL https://rupress.org/jcb/article/162/4/543/33337/Impaired-membrane-resealing-and-autoimmune.

[369] Ogata, A. *et al.* The role of angiopoietin-like protein 2 in pathogenesis of dermatomyositis. *Biochemical and Biophysical Research Communications* **418**, 494–499 (2012). URL https://linkinghub.elsevier.com/retrieve/pii/S0006291X12000757.

[370] Strongwater, S. L., Dorovini-zis, K., Ball, R. D. & Schnitzer, T. J. A Murine Model of Polymyositis Induced by Coxsackievirus B1 (Tucson Strain). *Arthritis & Rheumatism* **27**, 433–442 (1984). URL http://doi.wiley.com/10.1002/art.1780270411.

[371] Morrison, T. E. *et al.* A Mouse Model of Chikungunya Virus–Induced Musculoskeletal Inflammatory Disease. *The American Journal of Pathology* **178**, 32–40 (2011). URL https://linkinghub.elsevier.com/retrieve/pii/S0002944010000647.

[372] Kojima, T. *et al.* Myosin-induced autoimmune polymyositis in the rat. *Journal of the Neurological Sciences* **151**, 141–148 (1997). URL https://linkinghub.elsevier.com/retrieve/pii/S0022510X97001482.

[373] Kohyama, K. & Matsumoto, Y. C-protein in the skeletal muscle induces severe autoimmune polymyositis in Lewis rats. *Journal of Neuroimmunology* **98**, 130–135 (1999). URL https://linkinghub.elsevier.com/retrieve/pii/S0165572899000879.

[374] Nakano, J. *et al.* Laminin-Induced Autoimmune Myositis in Rats. *Journal of Neuropathology & Experimental Neurology* **64**, 790–796 (2005). URL https://academic.oup.com/jnen/article-lookup/doi/10.1097/01.jnen.0000178851.76056.0b.

[375] Romão, V. C. *et al.* Sjögren's syndrome: state of the art on clinical practice guidelines. *RMD Open* **4**, e000789 (2018). URL https://rmdopen.bmj.com/lookup/doi/10.1136/rmdopen-2018-000789.

[376] Xiao, F. *et al.* Animal models of Sjögren's syndrome: an update. *Clinical and Experimental Rheumatology* **37 Suppl 118**, 209–216 (2019).

[377] Hoffman, R. W., Alspaugh, M. A., Waggie, K. S., Durham, J. B. & Walker, S. E. Sjögren's syndrome in MRL/l and MRL/n mice. *Arthritis & Rheumatism* **27**, 157–165 (1984). URL http://doi.wiley.com/10.1002/art.1780270206.

[378] Miyagawa, J. *et al.* Ultrastructural and immunocytochemical aspects of lymphocytic submandibulitis in the non-obese diabetic (NOD) mouse. *Virchows Archiv. B, Cell Pathology Including Molecular Pathology* **51**, 215–225 (1986).

[379] Asamoto, H. *et al.* Infiltration of Lymphocytes in Various Organs of the NOD (Non-obese Diabetic) Mouse. *Journal of the Japan Diabetes Society* **27**, 775–781 (1984).

[380] Hu, Y., Nakagawa, Y., Purushotham, K. R. & Humphreys-Beher, M. G. Functional changes in salivary glands of autoimmune disease-prone NOD mice. *The American Journal of Physiology* **263**, E607–614 (1992).

[381] Robinson, C. P. *et al.* A novel NOD-derived murine model of primary Sjögren's syndrome. *Arthritis and Rheumatism* **41**, 150–156 (1998).

[382] Nguyen, C. *et al.* Sjögren's syndrome-like disease of C57BL/6.NOD-Aec1 Aec2 mice: gender differences in keratoconjunctivitis sicca defined by a cross-over in the chromosome 3 Aec1 locus. *Scandinavian Journal of Immunology* **64**, 295–307 (2006).

[383] Haneji, N., Hamano, H., Yanagi, K. & Hayashi, Y. A new animal model for primary Sjögren's syndrome in NFS/sld mutant mice. *Journal of Immunology (Baltimore, Md.: 1950)* **153**, 2769–2777 (1994).

[384] Tsubata, R. *et al.* Autoimmune disease of exocrine organs in immunodeficient alymphoplasia mice: a spontaneous model for Sjören's syndrome. *European Journal of Immunology* **26**, 2742–2748 (1996). URL http://doi.wiley.com/10.1002/eji.1830261129.

[385] Li, H., Dai, M. & Zhuang, Y. A T cell intrinsic role of Id3 in a mouse model for primary Sjogren's syndrome. *Immunity* **21**, 551–560 (2004).

[386] Shim, G.-J. *et al.* Aromatase-deficient mice spontaneously develop a lymphoproliferative autoimmune disease resembling Sjogren's syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 12628–12633 (2004).

[387] Kuroda, N. *et al.* Development of autoimmunity against transcriptionally unrepressed target antigen in the thymus of Aire-deficient mice. *Journal of Immunology (Baltimore, Md.: 1950)* **174**, 1862–1870 (2005).

[388] Oak, J. S. *et al.* Sjögren's syndrome-like disease in mice with T cells lacking class 1A phosphoinositide-3-kinase. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 16882–16887 (2006).

[389] Peng, B. *et al.* Defective feedback regulation of NF-kappaB underlies Sjogren's syndrome in mice with mutated kappaB enhancers of the IkappaBalpha promoter. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 15193–15198 (2010).

[390] Husain-Krautter, S., Kramer, J. M., Li, W., Guo, B. & Rothstein, T. L. The osteopontin transgenic mouse is a new model for Sjögren's syndrome. *Clinical Immunology (Orlando, Fla.)* **157**, 30–42 (2015).

[391] Ishimaru, N. *et al.* Expression of the retinoblastoma protein RbAp48 in exocrine glands leads to Sjögren's syndrome-like autoimmune exocrinopathy. *The Journal of Experimental Medicine* **205**, 2915–2927 (2008).

[392] Limaye, A. *et al.* Targeted TNF- Overexpression Drives Salivary Gland Inflammation. *Journal of Dental Research* **98**, 713–719 (2019).

[393] Scofield, R. H., Asfa, S., Obeso, D., Jonsson, R. & Kurien, B. T. Immunization with short peptides from the 60-kDa Ro antigen recapitulates the serological and pathological findings as well as the salivary gland dysfunction of Sjogren's syndrome. *Journal of Immunology (Baltimore, Md.: 1950)* **175**, 8409–8414 (2005).

[394] Zheng, J. *et al.* B Cells Are Indispensable for a Novel Mouse Model of Primary Sjögren's Syndrome. *Frontiers in Immunology* **8**, 1384 (2017).

[395] Iizuka, M. *et al.* Pathogenic role of immune response to M3 muscarinic acetylcholine receptor in Sjögren's syndrome-like sialoadenitis. *Journal of Autoimmunity* **35**, 383–389 (2010).

[396] Nishimori, I. *et al.* Induction of experimental autoimmune sialoadenitis by immunization of PL/J mice with carbonic anhydrase II. *Journal of Immunology (Baltimore, Md.: 1950)* **154**, 4865–4873 (1995).

[397] Bombardieri, M. *et al.* Inducible tertiary lymphoid structures, autoimmunity, and exocrine dysfunction in a novel model of salivary gland inflammation in C57BL/6 mice. *Journal of Immunology (Baltimore, Md.: 1950)* **189**, 3767–3776 (2012).

[398] Nguyen, C. Q. *et al.* Pathogenic effect of interleukin-17A in induction of Sjögren's syndrome-like disease using adenovirus-mediated gene transfer. *Arthritis Research & Therapy* **12**, R220 (2010).

[399] Young, N. A. *et al.* A chimeric human-mouse model of Sjögren's syndrome. *Clinical Immunology (Orlando, Fla.)* **156**, 1–8 (2015).

[400] Bokhari, S. R. A., Zulfiqar, H. & Mansur, A. Sarcoidosis. In *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2020). URL `http://www.ncbi.nlm.nih.gov/books/NBK430687/`.

[401] Altmann, D. M. & Boyton, R. J. Models of sarcoidosis. *Drug Discovery Today: Disease Models* **3**, 21–25 (2006). URL https://linkinghub.elsevier.com/retrieve/pii/S1740675706000193.

[402] Belcher, R. W. & Reid, J. D. Sarcoid granulomas in CBA/J mice. Histologic response after inoculation with sarcoid and nonsarcoid tissue homogenates. *Archives of Pathology* **99**, 283–285 (1975).

[403] Grizzanti, J. N. & Rosenstreich, D. L. Effect of inoculation of sarcoid tissue into athymic (nude) mice. *Sarcoidosis* **5**, 136–141 (1988).

[404] Iwai, K. & Takahashi, S. TRANSMISSIBILITY OF SARCOID-SPECIFIC GRANULOMAS IN THE FOOTPADS OF MICE. *Annals of the New York Academy of Sciences* **278**, 249–259 (1976). URL http://doi.wiley.com/10.1111/j.1749-6632.1976.tb47036.x.

[405] Soler, P., Bernaudin, J. F. & Basset, F. Ultrastructure of pulmonary granulomatosis induced in rats by intravenous complete Freund's adjuvant. *Virchows Archiv A Pathological Anatomy and Histology* **368**, 35–50 (1975). URL http://link.springer.com/10.1007/BF00432165.

[406] Bergeron, A. *et al.* Computed tomography of pulmonary sarcoid-like granulomas induced by complete Freund's adjuvant in rats. *European Respiratory Journal* **18**, 357–361 (2001). Publisher: Eur Respiratory Soc.

[407] Hu, Y., Yibrehu, B., Zabini, D. & Kuebler, W. M. Animal models of sarcoidosis. *Cell and Tissue Research* **367**, 651–661 (2017). URL http://link.springer.com/10.1007/s00441-016-2526-3.

[408] Besnard, V. & Jeny, F. Models Contribution to the Understanding of Sarcoidosis Pathogenesis: "Are There Good Models of Sarcoidosis?". *Journal of Clinical Medicine* **9**, 2445 (2020). URL https://www.mdpi.com/2077-0383/9/8/2445.

[409] Nishiwaki, T. *et al.* Indigenous Pulmonary Propionibacterium acnes Primes the Host in the Development of Sarcoid-Like Pulmonary Granulomatosis in Mice. *The American Journal of Pathology* **165**, 631–639 (2004). URL https://linkinghub.elsevier.com/retrieve/pii/S0002944010633275.

[410] Kishi, J. *et al.* Blockade of Th1 chemokine receptors ameliorates pulmonary granulomatosis in mice. *European Respiratory Journal* **38**, 415–424 (2011). URL http://erj.ersjournals.com/cgi/doi/10.1183/09031936.00070610.

[411] Werner, J. L. *et al.* Induction of Pulmonary Granuloma Formation by Propionibacterium acnes Is Regulated by MyD88 and Nox2. *American Journal of Respiratory Cell and Molecular Biology* **56**, 121–130 (2017).

[412] Minami, J. *et al.* Pulmonary granulomas caused experimentally in mice by a recombinant trigger-factor protein of Propionibacterium acnes. *Journal of Medical and Dental Sciences* **50**, 265–274 (2003).

[413] Chen, E. S. *et al.* Serum Amyloid A Regulates Granulomatous Inflammation in Sarcoidosis through Toll-like Receptor-2. *American Journal of Respiratory and Critical Care Medicine* **181**, 360–373 (2010). URL http://www.atsjournals.org/doi/abs/10.1164/rccm.200905-0696OC.

[414] Samokhin, A. O., Bühling, F., Theissig, F. & Brömme, D. ApoE-Deficient Mice on Cholate-Containing High-Fat Diet Reveal a Pathology Similar to Lung Sarcoidosis. *The American Journal of Pathology* **176**, 1148–1156 (2010). URL https://linkinghub.elsevier.com/retrieve/pii/S0002944010604294.

[415] Linke, M. *et al.* Chronic signaling via the metabolic checkpoint kinase mTORC1 induces macrophage granuloma formation and marks sarcoidosis progression. *Nature Immunology* **18**, 293–302 (2017). URL http://www.nature.com/articles/ni.3655.

[416] Locke, L. W., Schlesinger, L. S. & Crouser, E. D. Current Sarcoidosis Models and the Importance of Focusing on the Granuloma. *Frontiers in Immunology* **11**, 1719 (2020). URL https://www.frontiersin.org/article/10.3389/fimmu.2020.01719/full.

[417] Reijererk, E. P. R., Veldhuis Kroeze, E. J. B. & Sloet van Oldruitenborgh-Oosterbaan, M. Equine sarcoidosis. *Sarcoidosis, vasculitis, and diffuse lung diseases: official journal of WASOG* **26**, 20–23 (2009).

[418] Sloet van Oldruitenborgh-Oosterbaan, M. M. & Grinwis, G. C. M. Equine sarcoidosis: clinical signs, diagnosis, treatment and outcome of 22 cases: Equine sarcoidosis. *Veterinary Dermatology* **24**, 218–e48 (2013). URL http://doi.wiley.com/10.1111/j.1365-3164.2012.01108.x.

[419] Swaisgood, C. M. *et al.* Development of a Sarcoidosis Murine Lung Granuloma Model Using *Mycobacterium* Superoxide Dismutase A Peptide. *American Journal of Respiratory Cell and Molecular Biology* **44**, 166–174 (2011). URL http://www.atsjournals.org/doi/abs/10.1165/rcmb.2009-0350OC.

[420] Herndon, B., Quinn, T., Wasson, N., Nzabi, M. & Molteni, A. Urease and Helicobacter spp. Antigens in Pulmonary Granuloma. *Journal of Comparative Pathology* **148**, 266–277 (2013). URL https://linkinghub.elsevier.com/retrieve/pii/S0021997512001259.

[421] Huizar, I. *et al.* Novel Murine Model of Chronic Granulomatous Lung Inflammation Elicited by Carbon Nanotubes. *American Journal of Respiratory Cell and Molecular Biology* **45**, 858–866 (2011). URL http://www.atsjournals.org/doi/abs/10.1165/rcmb.2010-0401OC.

[422] Hill, A. & Hill, Q. A. Autoimmune hemolytic anemia. *Hematology. American Society of Hematology. Education Program* **2018**, 382–389 (2018).

[423] Dhaliwal, G., Cornett, P. A. & Tierney, L. M. Hemolytic anemia. *American Family Physician* **69**, 2599–2606 (2004).

[424] Howie, H. L. & Hudson, K. E. Murine models of autoimmune hemolytic anemia:. *Current Opinion in Hematology* **25**, 473–481 (2018). URL `http://journals.lww.com/00062752-201811000-00010`.

[425] Bérubé, N. G. *et al.* The chromatin-remodeling protein ATRX is critical for neuronal survival during corticogenesis. *The Journal of Clinical Investigation* **115**, 258–267 (2005).

[426] Medina, C. F. *et al.* Altered visual function and interneuron survival in Atrx knockout mice: inference for the human syndrome. *Human Molecular Genetics* **18**, 966–977 (2009).

[427] Merkle, S. & Pretsch, W. Glucose-6-phosphate isomerase deficiency associated with nonspherocytic hemolytic anemia in the mouse: an animal model for the human disease. *Blood* **81**, 206–213 (1993).

[428] Peters, L. L. *et al.* Purkinje cell degeneration associated with erythroid ankyrin deficiency in nb/nb mice. *The Journal of Cell Biology* **114**, 1233–1241 (1991).

[429] Rank, G. *et al.* Novel roles for erythroid Ankyrin-1 revealed through an ENU-induced null mouse mutant. *Blood* **113**, 3352–3362 (2009).

[430] Hughes, M. R. *et al.* A novel ENU-generated truncation mutation lacking the spectrin-binding and C-terminal regulatory domains of Ank1 models severe hemolytic hereditary spherocytosis. *Experimental Hematology* **39**, 305–320, 320.e1–2 (2011).

[431] Robledo, R. F. *et al.* Analysis of novel sph (spherocytosis) alleles in mice reveals allele-specific loss of band 3 and adducin in alpha-spectrin-deficient red cells. *Blood* **115**, 1804–1814 (2010).

[432] Unger, A. E., Harris, M. J., Bernstein, S. E., Falcone, J. C. & Lux, S. E. Hemolytic anemia in the mouse. Report of a new mutation and clarification of its genetics. *The Journal of Heredity* **74**, 88–92 (1983).

[433] Peters, L. L. *et al.* Identification of quantitative trait loci that modify the severity of hereditary spherocytosis in wan, a new mouse model of band-3 deficiency. *Blood* **103**, 3233–3240 (2004).

[434] Kanno, H. *et al.* Primary structure of murine red blood cell-type pyruvate kinase (PK) and molecular characterization of PK deficiency identified in the CBA strain. *Blood* **86**, 3205–3210 (1995).

[435] Min-Oo, G. *et al.* Pyruvate kinase deficiency in mice protects against malaria. *Nature Genetics* **35**, 357–362 (2003).

[436] Kautz, L. *et al.* Identification of erythroferrone as an erythroid regulator of iron metabolism. *Nature Genetics* **46**, 678–684 (2014).

[437] Chen, M. L. *et al.* Erythroid dysplasia, megaloblastic anemia, and impaired lymphopoiesis arising from mitochondrial dysfunction. *Blood* **114**, 4045–4053 (2009).

[438] Kong, Y. *et al.* Loss of alpha-hemoglobin-stabilizing protein impairs erythropoiesis and exacerbates beta-thalassemia. *The Journal of Clinical Investigation* **114**, 1457–1466 (2004).

[439] Jamsai, D. *et al.* A humanized mouse model for a common beta0-thalassemia mutation. *Genomics* **85**, 453–461 (2005).

[440] Yang, B. *et al.* A mouse model for beta 0-thalassemia. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 11608–11612 (1995).

[441] Nai, A. *et al.* Deletion of TMPRSS6 attenuates the phenotype in a mouse model of -thalassemia. *Blood* **119**, 5021–5029 (2012).

[442] Perkins, A. C., Sharpe, A. H. & Orkin, S. H. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* **375**, 318–322 (1995).

[443] Brown, F. C. *et al.* ENU mutagenesis identifies the first mouse mutants reproducing human -thalassemia at the genomic level. *Blood Cells, Molecules & Diseases* **50**, 86–92 (2013).

[444] Chui, D. *et al.* Alpha-mannosidase-II deficiency results in dyserythropoiesis and unveils an alternate pathway in oligosaccharide biosynthesis. *Cell* **90**, 157–167 (1997).

[445] Mei, Y. *et al.* Ineffective erythropoiesis caused by binucleated late-stage erythroblasts in mDia2 hematopoietic specific knockout mice. *Haematologica* **101**, e1–5 (2016).

[446] White, R. A. *et al.* Hematologic characterization and chromosomal localization of the novel dominantly inherited mouse hemolytic anemia, neonatal anemia (Nan). *Blood Cells, Molecules & Diseases* **43**, 141–148 (2009).

[447] Peters, L. L. *et al.* Mild spherocytosis and altered red cell ion transport in protein 4. 2-null mice. *The Journal of Clinical Investigation* **103**, 1527–1537 (1999).

[448] Gilligan, D. M. *et al.* Targeted disruption of the beta adducin gene (Add2) causes red blood cell spherocytosis in mice. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 10717–10722 (1999).

[449] Patterson, A. D., Hollander, M. C., Miller, G. F. & Fornace, A. J. Gadd34 requirement for normal hemoglobin synthesis. *Molecular and Cellular Biology* **26**, 1644–1653 (2006).

[450] Popp, R. A. & Enlow, M. K. Radiation-induced alpha-thalassemia in mice. *American Journal of Veterinary Research* **38**, 569–572 (1977).

[451] Loyd, M. R., Okamoto, Y., Randall, M. S. & Ney, P. A. Role of AP1/NFE2 binding sites in endogenous alpha-globin gene transcription. *Blood* **102**, 4223–4228 (2003).

[452] Bender, M. A., Bulger, M., Close, J. & Groudine, M. Beta-globin gene switching and DNase I sensitivity of the endogenous beta-globin locus in mice do not require the locus control region. *Molecular Cell* **5**, 387–393 (2000).

[453] Shehee, W. R., Oliver, P. & Smithies, O. Lethal thalassemia after insertional disruption of the mouse major adult beta-globin gene. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 3177–3181 (1993).

[454] Lewis, J. *et al.* A common human beta globin splicing mutation modeled in mice. *Blood* **91**, 2152–2156 (1998).

[455] Fiering, S. *et al.* Targeted deletion of 5'HS2 of the murine beta-globin LCR reveals that it is not essential for proper regulation of the beta-globin locus. *Genes & Development* **9**, 2203–2213 (1995).

[456] Huo, Y., McConnell, S. C. & Ryan, T. M. Preclinical transfusion-dependent humanized mouse model of beta thalassemia major. *Blood* **113**, 4763–4770 (2009).

[457] Kuiper, J. J. W. *et al.* Functionally distinct ERAP1 and ERAP2 are a hallmark of HLA-A29-(Birdshot) Uveitis. *Human Molecular Genetics* (2018). URL `https://academic.oup.com/hmg/advance-article/doi/10.1093/hmg/ddy319/5095324`.

[458] Kuiper, J. J. W. & Venema, W. J. HLA-A29 and Birdshot Uveitis: Further Down the Rabbit Hole. *Frontiers in Immunology* **11**, 599558 (2020). URL `https://www.frontiersin.org/articles/10.3389/fimmu.2020.599558/full`.

[459] Vallejo-Díaz, J., Chagoyen, M., Olazabal-Morán, M., González-García, A. & Carrera, A. C. The Opposing Roles of PIK3R1/p85 and PIK3R2/p85 in Cancer. *Trends in Cancer* **5**, 233–244 (2019).

[460] Nunes-Santos, C. J., Uzel, G. & Rosenzweig, S. D. PI3K pathway defects leading to immunodeficiency and immune dysregulation. *The Journal of Allergy and Clinical Immunology* **143**, 1676–1687 (2019).

[461] de Jesus, A. A. & Goldbach-Mansky, R. Newly recognized Mendelian disorders with rheumatic manifestations. *Current Opinion in Rheumatology* **27**, 511–519 (2015).

[462] Jamee, M. *et al.* Clinical, Immunological, and Genetic Features in Patients with Activated PI3K Syndrome (APDS): a Systematic Review. *Clinical Reviews in Allergy & Immunology* **59**, 323–333 (2020).

[463] Mont'Alverne, A. R. S. *et al.* Diminished ovarian reserve in Behçet's disease patients. *Clinical Rheumatology* **34**, 179–183 (2015). URL `http://link.springer.com/10.1007/s10067-014-2680-5`.

[464] Cil, A. P., Karabulut, A. A. & Koçak, M. Assessment of ovarian stromal artery Doppler characteristics and serum hormone levels in patients with Behçet disease. *Diagnostic and Interventional Radiology (Ankara, Turkey)* **16**, 288–292 (2010).

[465] Lin, X. *et al.* IL-10-producing regulatory B cells restrain the T follicular helper cell response in primary Sjögren's syndrome. *Cellular & Molecular Immunology* **16**, 921–931 (2019). URL http://www.nature.com/articles/s41423-019-0227-z.

[466] Howie, J. & Helyer, B. The Immunology and Pathology of NZB Mice. In *Advances in Immunology*, vol. 9, 215–266 (Elsevier, 1968). URL https://linkinghub.elsevier.com/retrieve/pii/S0065277608604447.

[467] Playfair, J. H. L. & Marshall-Clarke, S. Induction of Red Cell Autoantibodies in Normal Mice. *Nature New Biology* **243**, 213–214 (1973). URL http://www.nature.com/articles/newbio243213a0.

[468] Okamoto, M. *et al.* A transgenic model of autoimmune hemolytic anemia. *The Journal of Experimental Medicine* **175**, 71–79 (1992).

[469] Huo, Y. *et al.* Humanized Mouse Model of Cooley's Anemia. *The Journal of Biological Chemistry* **284**, 4889–4896 (2009).