Fall 2020

CAN VIRUSES CAUSE LUPUS?

Stephen I. Reeder
Southeastern University - Lakeland

CAN VIRUSES CAUSE LUPUS?

by

Stephen Isaac Reeder

Submitted to the School of Honors Committee

in partial fulfillment

of the requirements of University Honors Scholars

Southeastern University

2020

Dedication

I would like to dedicate this paper to Ms. Angela, who served as my inspiration for choosing lupus as my topic for my thesis. She is an incredible woman whose body wages a war with her and yet she still manages to fight and live her life. She has raised an incredible young man who is my good friend, and she lives her life with an unwavering commitment to God.

Acknowledgments

I would like to thank first Dr. Schraw who served as my advisor for my thesis. His guidance helped me to keep focused on the things that mattered when constructing my thesis, and his knowledge helped direct me to ideas that served as the foundation for my thesis. I would also like to thank my future wife, Abi, for her love and support as I have worked through the process of writing my thesis. I also want to thank my family for supporting me my whole life and shaping me into the person that I am today. Finally, I want to thank God for being my source of strength and life. Without God I am nothing. The gifts and talents he has given me, the people he has placed in my life, and the strength to complete something as extensive as this thesis come from him, and I wouldn't be who I am and where I am today without his mercy and his grace.

Abstract

Like so many autoimmune diseases, the exact cause of systemic lupus erythematosus (SLE) remains unknown. Evidence points to both genetics and environment playing roles in the onset of the disease, but neither acts independent of the other. Genetics are the easier of the two to study with recent advances in the field making it easier to isolate genes shared by individuals with the disease. However, genetic studies reveal that there is almost certainly an environmental component to the development of SLE. The underlying pathology and existing research on environmental contributors to the development of SLE suggest that viruses could potentially be an environmental factor that leads to the onset of SLE. Research has been done in the past in an attempt to establish a connection between viruses and the onset of SLE; however, these studies have been limited to providing circumstantial evidence due to the limits of existing technologies. A recent technology called VirScan developed by researchers at Harvard holds the potential to overcome the limitations of past research. The purpose of this thesis is to present an experimental approach to use VirScan to determine which viruses leading to the development of lupus.

KEY WORDS: lupus, SLE, virus, VirScan, autoimmune disease, autoantibodies

Table of Contents

Introduction....1

Methodology....3

Literature Review

Systemic Lupus Erythematosus....4

Overview of Disease....4

Antibodies & Autoantibodies....6

Apoptosis....12

Inflammation....15

Genetic & Environmental Factors....18

Viruses & Autoimmunity....21

Viruses....21

Human Immune Response to Viruses....23

Molecular Mimicry....27

Bystander Activation....31

Epitope Spreading....32

Connecting Viruses to Autoimmune Diseases....34

VirScan....36

High Throughput DNA Sequencing....36

De Novo DNA Synthesis....38

T7 Phage Display System....40

Phage Immunoprecipitation Sequencing (PhIP-seq)....41

VirScan Design....43

Conclusion....48

References....51

Introduction

For the average person, the fatigue and achiness associated with a bad flu is the only experience that can quantify a day in the life of someone living with SLE.[1] Except that chronic, widespread pain and tiredness is experienced every day.[1] For people living with SLE, something as simple as getting out of bed becomes a difficult and painful task.[1] What's worse is there is no cure, only symptom management that makes the pain bearable.[2] Like a thief, SLE comes and steals some of the basic pleasures of life that can so easily be taken for granted like being able go through the day not feeling tired all the time or being able to perform everyday tasks without pain. Where the thief comes from is unknown and how to keep it from stealing from its victims has proven difficult to prevent. Most people are unaware of the battle that people with lupus face on a daily basis and these people often go unnoticed.[1] SLE takes something different from each person, with people often experiencing a broad array of symptoms with numerous combinations of these symptoms.[2]

SLE is a difficult disease to quantify because the symptoms vary so greatly between individuals.[2] Tiredness and chronic widespread pain are only a couple of the many different symptoms that people with SLE experience.[2] SLE often goes undiagnosed and is not a disease that is reportable so estimates of the number of people who have SLE are difficult to obtain.[3] The Lupus Foundation of America has attempted to compile several statistics to help quantify the existing prevalence and impact of the four forms of lupus, among which SLE is the most prevalent.[4] They estimate that as many as 1.5 million Americans live with lupus, and at least 5 million people worldwide, have one of the four forms of lupus.[4] The demographic most heavily affected by lupus are women of childbearing age, with 9 out of 10 people with lupus being women between the ages of 15 and 44.[4] Research has also shown that there is a significant

financial impact on individuals living with lupus.[5] On average, the cost of medical care directly related to the treatment of lupus is $33,223 a year, and the financial losses associated with limitations to productivity caused by lupus range from $1,252 and $20,046 annually.[5] This can add up to a total cost of $50,000 a year in expenses related to lupus. Studies have also shown that the severity of lupus is worse and mortality rates are higher among minority groups.[6] Women of color are 2 to 3 times more likely to develop lupus compared to Caucasian women.[7] Among the few positives associated with the disease is a low mortality rate with only about 10-15% of people with lupus suffering from a premature death.[7]

Like so many other autoimmune diseases, treatment for SLE is mainly focused on managing symptoms and the cause of the disease is unknown.[8] Evidence points to the cause of SLE being multifaceted, but there is almost certainly a genetic and an environmental component to the development of lupus.[8] The environmental component offers a potential target for the prevention of SLE, and the investigation of this avenue will be the focus of this thesis. Investigating an environmental cause of SLE requires a thorough examination of the disease pathology and pathogenesis. In conjunction with this, the potential mechanisms of interaction between the environmental agent must be well understood and possess a certain level of feasibility that makes further investigation justifiable. Finally, it must then be possible to investigate this connection with a technology that can prove the connection between the environmental factor and the development of SLE. For the purpose of this thesis, each of these elements will be investigated in detail and the environmental factor in question will be viruses. All of these components will provide a wholistic picture of the available data that seeks to answer the question, “Can viruses cause SLE?” If so, “Can the viruses causing SLE be stopped, leading to the prevention of SLE?”

Methodology

The methodology used to construct this thesis was an extended review of available literature pertaining to the relationship between viruses and the development of lupus, specifically systemic lupus erythematosus (SLE). To collect data pertaining to this topic, databases such as PubMed, PMC, and SciFinder were used to collect relevant and recent data on the topics investigated. The primary sources used in this paper were articles published in major scientific journals such as *Nature*, *Science*, and *Frontiers in Immunology*. The topics investigated include SLE, viruses and autoimmunity, and VirScan. Each of these topics were broken down into their key components and explained in detail for the purpose of synthesizing a thesis that illustrates the connection between viruses and SLE, and also introduces a technology that could be used to then investigate the connection between the two.

By first investigating the disease pathology and pathogenesis of SLE, the groundwork is laid to show how the pathogenesis could correlate with a viral infection. Then, by further investigating possible mechanisms, the theory is given feasibility as well as credibility for being a potential causative agent of SLE. Finally, by introducing a technology that could be employed to test this theory, this thesis completes the case for future investigation of this connection that has been thoroughly supported by existing literature throughout the course of this paper. To justify extensively researching a viral cause for lupus, there must be compelling evidence that supports this hypothesis. The methodology of this thesis seeks to present the evidence necessary to justify future research with an extended review of available literature on relevant topics.

Systemic Lupus Erythematosus

Overview of Disease

Systemic Lupus Erythematosus (SLE) is the disease most commonly associated with the term lupus, however it is one among four similar disease processes: neonatal and pediatric lupus erythematosus (NLE), discoid lupus erythematosus (DLE), drug-induced lupus (DIL), and systemic lupus erythematosus (SLE).[2] It is the most common and widely studied of the four types of lupus, which is why the term lupus is often used to refer to SLE.[2] The word "lupus" means wolf in Latin and became associated with SLE due to early descriptions of the disease dating back to the Middle Ages.[9] One of the common clinical manifestations of the disease is a malar rash that can have an appearance similar to that of a wolf bite, which is where the disease derives its association with the Latin word for wolf.[9] The clinical title, systemic lupus erythematosus, acquires most of its name from the disease manifestations relating to the skin.[2] The word erythematosus comes from the Greek word for redness, "erythros", referring to the redness observed on the skin and lupus as stated previously also relates to the skin; however, it is the word systemic that perhaps is the most accurate in describing the nature of this disease.[2] SLE affects almost every organ system in the body, among which it most commonly affects the skin, joints, and kidneys.[2]

As a consequence of the many ways SLE can affect the body, for someone to receive a diagnosis for SLE they must meet at least four of the eleven criteria based on the most common symptoms.[10] Three of the eleven symptoms used to identify SLE are related to the skin.[10] The first of which is a malar rash (Figure 1), also called the "butterfly rash" due to its location on the bridge of

Figure 1. Malar rash.[11]

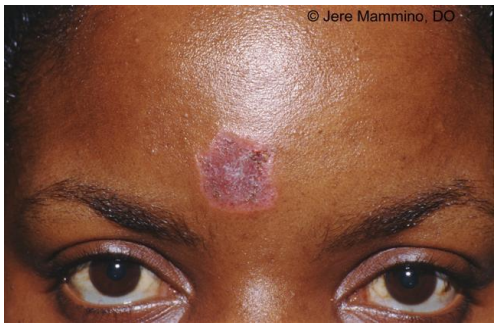

Figure 2. Discoid rash.[12]

the nose and the cheeks.[10,11] The second manifestation of SLE on the skin is a discoid rash (Figure 2) which is a more severe type of rash that can lead to scarring, and the final criteria related to the skin is a general sensitivity to sun exposure.[10,12] Lupus can also manifest within the body as ulcers on the mucosa of the nose and mouth (Figure 3), and as inflammation of the serosal membranes.[10,13] The sixth criteria for diagnosing SLE is inflammation of the joints, which is among the most common symptoms of SLE; however, like any of the eleven criteria it is not specific enough to diagnose SLE by itself.[10] Another criteria that is among the most common complications of SLE is renal disorders such as abnormal urine protein and diffuse proliferative glomerulonephritis.[10] Conversely, a more rare complication is the development neurological disorders such as seizures and psychosis.[10] The ninth criteria for SLE diagnosis is a number of hematologic disorders such as anemia, thrombocytopenia and leukopenia.[10] Finally, the last two criteria for the diagnosis of SLE involve the presences of autoantibodies, which are among the most definitive indicators of SLE.[10] Autoantibodies point to the reason why SLE is classified as an autoimmune disease.[10] The specific criteria for diagnosing lupus differentiate antinuclear antibodies as a separate criterion from the final criteria, which is any other autoantibodies such as anti-smith, anti-dsDNA and anti-phospholipid antibodies.[10]

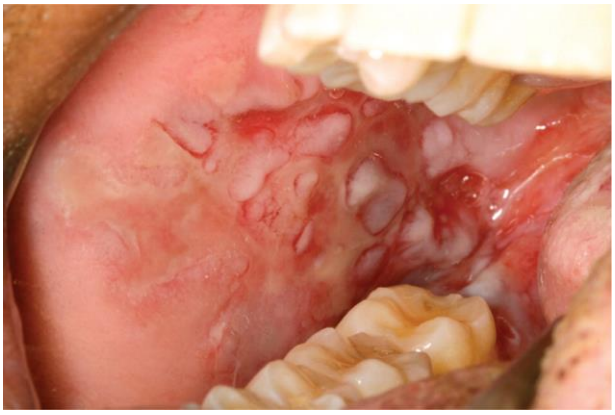

Figure 3. Ulcers on the oral mucosa.[13]

None of these eleven criteria are enough to diagnose lupus by themselves since each one of these criteria can be implicated in other disease processes.[10] However, when four or more of these symptoms are present, they are an indicator of a greater underlying problem.[14] Every person who deals with SLE experiences a different array of symptoms with a great deal of variety in the number of symptoms as well as the severity of these symptoms.[14] Most people experience periods of flare-ups where the symptoms become worse, and it is these flare-ups that most drug treatments are designed to help prevent or reduce.[14] Each element of SLE points back to the greater underlying issue of the immune system's failure to function properly.[15] An autoimmune disease is characterized by one's own immune system attacking self.[16] Autoantibodies, one of the most specific indicators of SLE, are evidence of the body's attempts to attack self instead of foreign pathogens. There are several different elements besides autoantibodies that culminate in the disease process of SLE such as apoptosis, inflammation, genetic factors and environmental factors.[16] Each one of these elements adds to the complexity and diversity of this disease, but each points back immune system's inability to recognize self and the aberrant responses that eventually lead to the destruction of self.

Antibodies & Autoantibodies

Preceding an understanding of why the body attacks itself in autoimmune diseases such as SLE, is foundational knowledge of how the body normally attacks foreign pathogens. Among the ways that immune system attacks foreign pathogens is the generation of highly variable proteins called immunoglobulins (Ig) or more commonly, antibodies.[17] These types of antibodies are shaped like the letter "Y" and have two identical regions on each side that have very specific binding capabilities (Figure 4).[17] These regions on the antibody called the paratopes bind to a

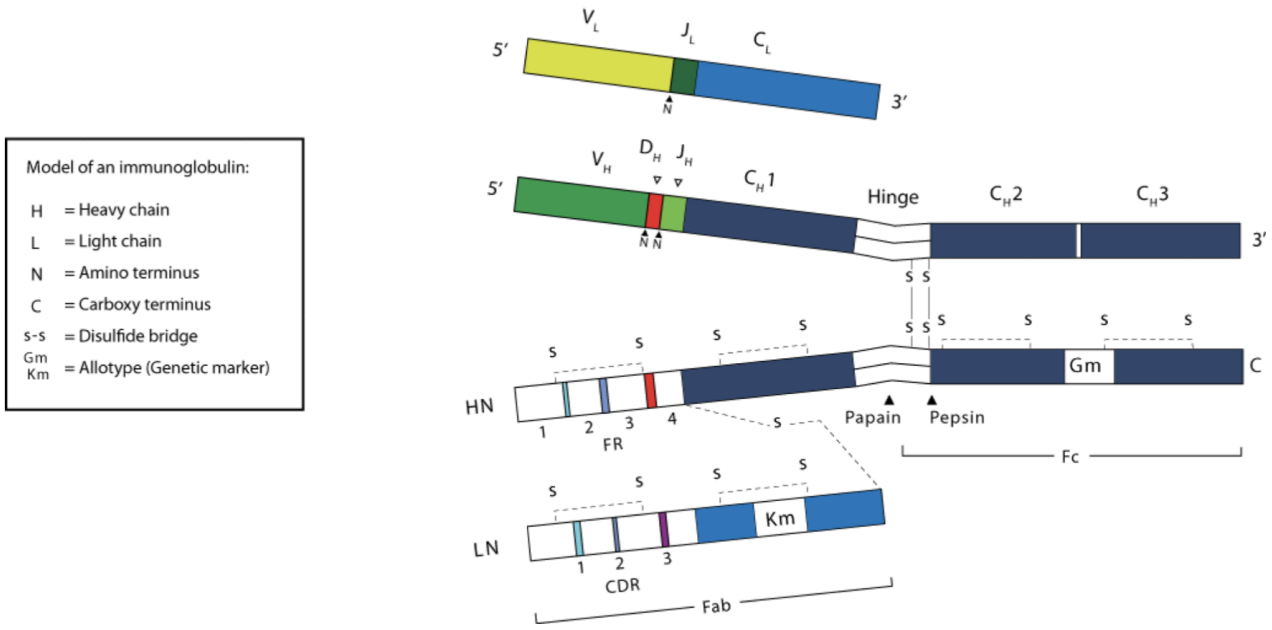

Figure 4. Two-dimensional model of an IgG molecule.[17]

corresponding region called the epitope on an antigen.[17] An antigen is the specific target that an antibody binds to and can be a number of different things including bacteria, viruses, cells, proteins, and foreign particles.[17] Normally, the body produces antibodies that are specific to foreign invaders such as viruses and bacteria, but it is also possible for the body to produce antibodies that are specific to molecules that are normally found in the body.[17]

Antibodies are produced by cells in the immune system called B lymphocytes or B cells.[17] There are five isotypes of immunoglobulins (IgM, IgG, IgA, IgE and IgD) (Figure 5), and each possesses a set of functions and targets.[17] The preliminary form in which antibodies are presented is as IgM molecules, which exist as a monomers on the surface of B cells or as a pentamers in serum.[17] The role of IgM is to be less specific and serve as a first line of attack against new antigens.[18] On the surface of B cells, IgM serves as a screener for immature B cells that have yet to differentiate.[18] Different immature B cells present specific IgM monomers on their surface with low affinity and high reactivity for the purpose of identifying new targets and

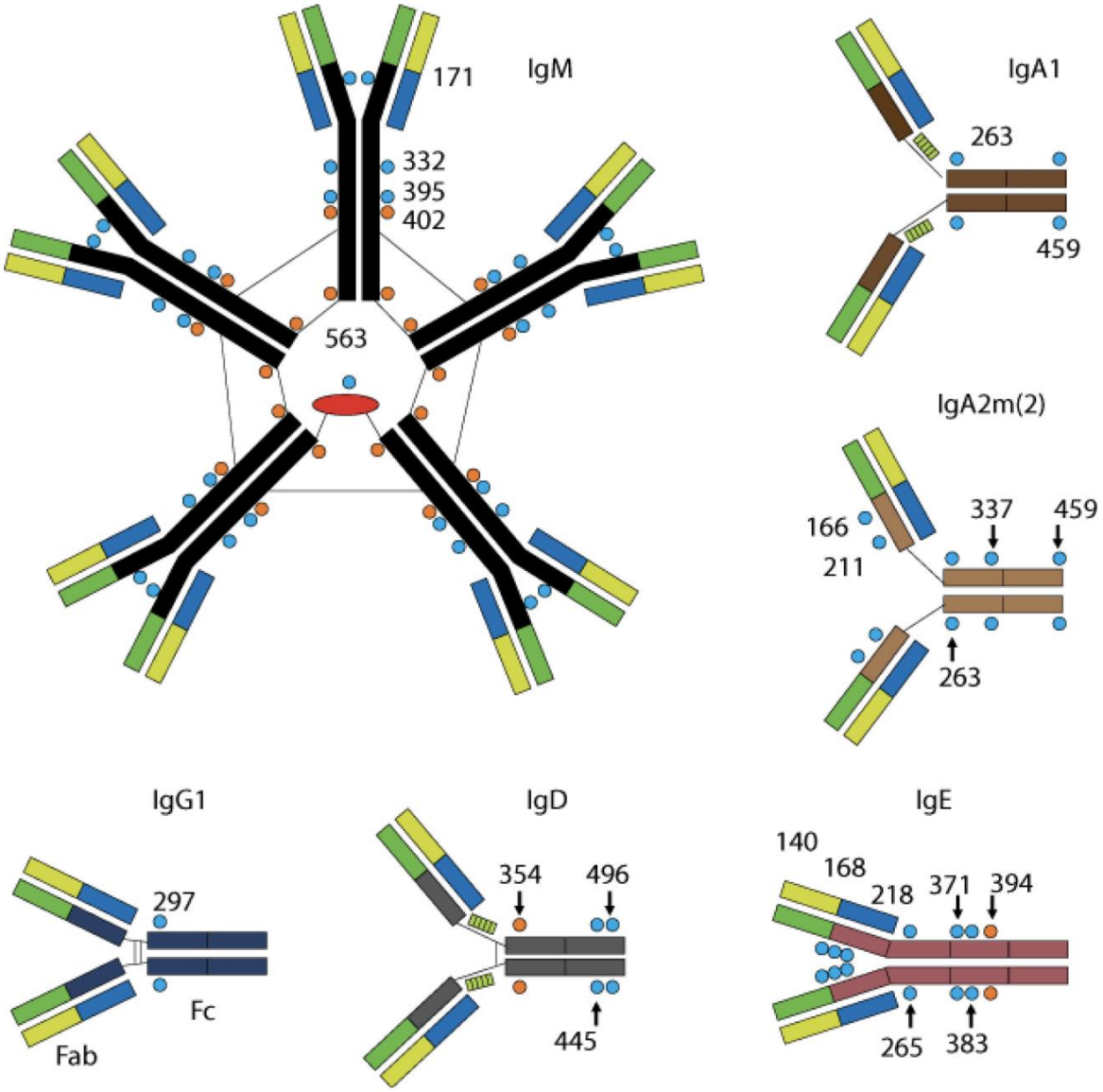

Figure 5. Two-dimensional models of immunoglobulin sturctures.[17]

activating a more specific response.[18] IgM antibodies are also secreted in serum as pentamers to act as an opsonins.[18] An opsonin is a type of biochemical marker that labels antigens for destruction by phagocytes and also serves as an adhesive that makes antigens, such as bacteria, easier to phagocytize.[18] IgM is not the only class of antibody that can act as an opsonin, and in their role as an opsonin, antibodies serve as activators for the complement system, which will be described in more detail later.[17] The most common type of immunoglobulin, IgG, mainly serves

as an opsonin.[19] This isotype is secreted by differentiated B cells called plasma cells, and acts as the central component to the immune system's ability to respond effectively to antigens to which it has been previously exposed.[19] Aside from floating in serum acting as an opsonin and activating the complement system, IgG molecules bind to FcγR receptors on T-cells which activates the response of T-cells.[19] IgG executes the most common function associated with antibodies, however the other three classes of antibodies, while less abundant, serve specific and vital roles in normal immune function.[17] IgA molecules are found in the mucosal linings and serve to activate the response of mucosal tissues to foreign invaders.[17] The presence of IgA antibodies stimulates mucus production, promotes inflammatory response via the complement system and acts as an opsonin on antigens present on the surface of the mucosal lining.[17] IgE immunoglobulins specialize in activating granulocytes such as mast cells, basophils, Langerhans cells and eosinophils.[17] These cells possess FcεRI receptors which have an extremely high affinity for IgE molecules and activate the immune response generally associated with allergies, which is designed to target larger pathogens such as parasitic worms.[17] The last isotype, IgD has the smallest and least understood role, which is believed to be involved in B cell maturation and serves as a B cell receptor (BCR) alongside IgM.[17]

Normally, antibodies do not have a high affinity for molecules regularly found in the body, but in people with SLE there is a high prevalence of these kinds of antibodies.[15] These self-reacting antibodies are called autoantibodies.[15] In healthy individuals, these autoantibodies or natural antibodies are present in low amounts as mainly IgM molecules.[15] It is believed that these self-reacting IgM antibodies play a maintenance role in healthy individuals and have the capacity to become harmful when perpetuated.[15] The autoantibodies become dangerous when they class switch from IgM to IgG molecules.[15] As IgM molecules, autoantibodies react more

broadly and are less adhesive, but if mutations occur and the autoantibodies switch classes to IgG, they become more specific and widespread.[15] Normally, the body has mechanisms in place to prevent this process from taking place, by selecting for B cells that are not self-reactive through a series of checkpoints.[20] If this process is not effective in eliminating self-reactive B cells, the widespread production of IgG autoantibodies is what eventually leads to the pathogenic role of autoantibodies.[15]

There are a number of self-antigens that can be the target of autoantibodies, with some being more common than others and the specificity of the autoantibodies potentially offering clues to how SLE will affect the body.[20] One classification of autoantibodies, Anti-Nuclear Antibodies (ANAs), is among the eleven criteria used to diagnose SLE, however it is not specific to SLE, so it cannot considered conclusive evidence of SLE.[20] ANA is a junk drawer term used to describe internal components of the cell, and it is not an entirely accurate term since it would seem to imply that it refers only to components contained within the nucleus.[20] The term also includes other components of the cell such as the mitotic spindle apparatus and cytoplasmic organelles.[20] Typically these components are contained within the cell and would not present any problems even in the presence of ANAs, but in some cases, such as after apoptosis, these components can be found outside of the cell, exposing them to ANAs if present.[20] The nature of how components normally contained within the cell become exposed to autoantibodies, eludes to the greater role that apoptosis plays in the disease process of SLE.[20] Double stranded DNA is another molecule that patients with SLE have been shown to develop antibodies against, with 60-90% of individuals with SLE presenting anti-dsDNA antibodies.[20] The immune system becomes exposed to dsDNA in the form of histones, and anti-dsDNA can be specific to different structures of DNA including the elongated nucleosome linker B-DNA (double-helix, right

handed turn), phosphodeoxiribose backbones, higher-order bent DNA structure, Z-DNA, and cruciform DNA structures (left handed turn).[20] The presence of anti-dsDNA is also commonly associated with lupus nephritis (LN) and the renal complications that occur with SLE.[20] Similar to anti-dsDNA antibodies, but not used as a diagnostic marker for SLE, many individuals with SLE also produce anti-nucleosome antibodies.[20] The nucleosome is the larger structure within which DNA is contained, and the presence of anti-nucleosome antibodies is an even more consistent indicator SLE than anti-dsDNA, with many of its implications being the same.[20] Another group of antigens which SLE patients often produce antibodies against are Sm antigens, which are named after the patient, Stephanie Smith, in whom these antibodies were first discovered.[20] There are seven core proteins (B, D1, D2, D3, E, F, G) which form a ring for small nuclear ribonucleoproteins (snRNP), and it is these proteins against which anti-Sm antibodies are formed.[20] In addition to anti-Sm antibodies, individuals with SLE may also form antibodies against other snRNP including anti-RNP, anti-Ro/SSA and anti-La/SSB antibodies, however these antibodies are less common than anti-Sm and are not used as a diagnostic marker for SLE.[20] The final type of autoantibody that is used as a diagnostic marker for SLE is anti-phospholipid (aPL).[20] While found in 30-40% of SLE patients, aPLs are found in other autoimmune diseases, infections, drug induced disorders, and some healthy controls, which renders them an ineffective indicator of SLE without the presence of other autoantibodies which are more specific to SLE.[20] The presence of aPLs can often lead to the development of antiphospholipid syndrome, which is a disorder not specific to SLE that is characterized by recurrent arterial or venous thrombosis, pregnancy-related problems, thrombocytopenia, hemolytic anemia, and persistent elevated levels of aPLs.[20] Many other autoantibodies have been discovered in association with SLE including anti-C1q antibodies, anti-ribosomal P (anti-P)

antibodies, anti-NMDAR antibodies, anti-annexin antibodies and others which are being newly identified.[20] However, these are not common enough to be used as a diagnostic criteria for SLE and their role in disease progression is still largely unknown.[20] The great diversity of autoantibodies produced by individuals with SLE points to the larger problem of lack of self-tolerance in conjunction with an increased presence of internal cell components which are found outside of the cell as a result of apoptosis. When these two elements are viewed simultaneously, the underlying pathogenesis of SLE begins to take shape.

Apoptosis

When considering the prevalence of autoantibodies, it must also be taken into consideration if there is a process that contributes to the presence and pathogenicity of these autoantibodies. The role of apoptosis in patients with SLE has become increasingly implicated as a major contributor to the disease pathogenesis, and can be directly linked to the problematic nature of autoantibodies.[15] Apoptosis is the process by which cells undergo programmed self-destruction.[21] This is an important homeostatic function of the body that counters the process of mitosis, controlling the number of cells present in the body.[21] In healthy individuals, this process occurs regularly without an inflammatory effect, however in individuals with SLE, the breakdown of this process can lead to the extracellular exposure of intracellular components, and can lead to the production of autoantibodies.[22] Cells undergoing apoptosis begin the process by condensing the chromatin present in the nucleus and exhibiting an overall cell shrinkage (Figure 6).[21,23] It is the process of cell shrinkage that is the distinguishing marker of apoptosis and what differentiates it from necrosis which is the alternative form of cell death.[21] Necrosis is an uncontrolled process by which cells die from external stressors or cell swelling and subsequent rupture.[21] Conversely, apoptosis is a highly controlled process with a vast array of genetic

regulators, and it is the malfunction of these regulators that is believed to be a contributing factor to the development of SLE.[22] In addition to the failure of typical regulators of apoptosis, evidence also suggests that the irregular buildup of oxidative stressors may induce irregular apoptosis in individuals with SLE.[24] During the normal process of apoptosis, following cell shrinkage, the cell initiates a process call plasma membrane blebbing, where the reduced cell divides into smaller fragments which maintain the containment of intracellular components within the plasma membrane.[21] These smaller packages of the fragmented cell express signals on their surface that act as signals for destruction by phagocytes such as macrophages, parenchymal cells, or neoplastic cells.[21,22] The molecules expressed on the cell surface include phosphatidylserine (PS), phosphatidylcholine (PC), and phosphatidylethanolamine (PE), and in addition to these signals on the cell surface, the release of adenosine triphosphate (ATP) and uridine triphosphate (UTP) also recruits phagocytes to engulf the apoptotic bodies.[22] In conjunction with abnormal rates of apoptosis, there also appears to be a decreased ability to clear apoptotic bodies in patients with SLE.[22] If the apoptotic bodies are not properly cleared, the plasma membrane begins to breakdown, leading to what is called secondary necrosis.[22] Secondary necrosis results in the internal cell components being released into the extracellular space.[22]

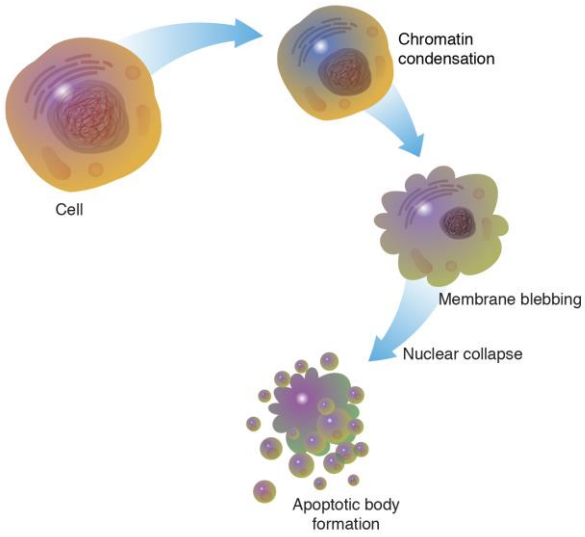

Figure 6. Stages of apoptosis.[23]

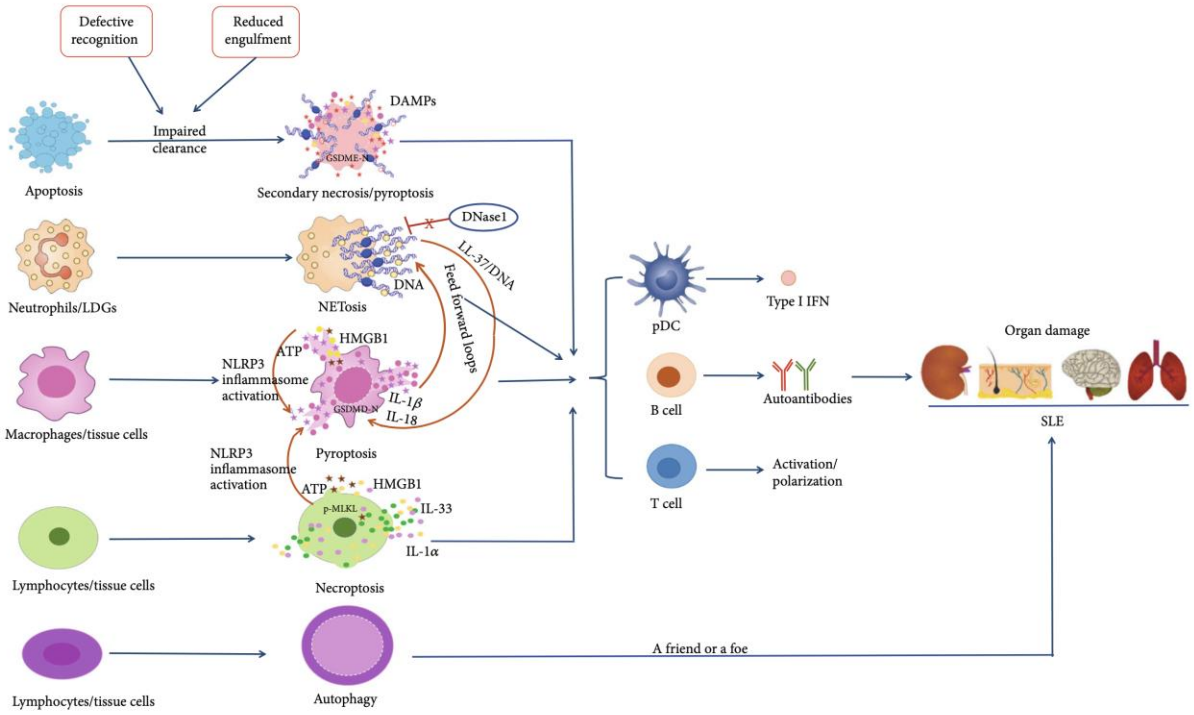

Figure 7. Programmed cell death pathways.[22]

In addition to typical routes of apoptosis being dysregulated and dysfunctional, there are also atypical routes of apoptosis that have been implicated in the pathogenesis of SLE (Figure 7).[22] The most likely contributor to SLE is a specific type of cell death that occurs in neutrophils called NETosis.[22] This form of cell death results in the formation of a neutrophil extracellular trap (NET), which is composed of decondensed chromatin decorated with intracellular components, including neutrophil elastase (NE), myeloperoxidase (MPO), high mobility group protein B1 (HMGB1), proteinase 3 (PR3), and LL-37.[22] Specifically the presence of decondensed chromatin seems to implicate NETosis as a likely contributor to the formation of SLE, since the presence of anti-dsDNA is among the main indicators of SLE.[22] Patients with SLE have also been shown to have higher levels of a specific type of neutrophil called low-density granulocytes (LDGs) which release higher amounts of inflammatory signals and show an

increased susceptibility to spontaneously undergo NETosis.[22] Compounding the effects of the increased activity of NETosis, SLE patients show a decreased ability to degrade NETs compared to healthy individuals and the increased presence of autoantibodies in SLE patients also hinders the breakdown of NETs.[22] Along with NETosis, other atypical routes of cell death including pyroptosis and necroptosis have been shown to contribute to disease manifestation of SLE, with both leading to intracellular components being exposed to the extracellular space.[22] The evidence points to a consistent failure in individuals with SLE to properly regulate cell death and clear cell debris following cell death, which then correlates with the increased presence of autoantibodies produced in response to the excess of intracellular materials present in the extracellular space. These factors combine to culminate in the excessive inflammatory response produced by the immune system, which is responsible for the many clinical disease manifestations of SLE.

Inflammation

At the junction of the previously described factors of dysregulated apoptosis and the subsequent prevalence of autoantibodies, the disease pathology of SLE begins to take shape. It is the nature of the class of diseases to which SLE belongs, autoimmune diseases, and its systemic nature that point to how these factors combine to produce widespread and constant pain for individuals with SLE. Autoimmune diseases are characterized by one's own immune system using mechanisms designed to combat foreign pathogens to attack self. The production of antibodies that bind to self-molecules leads to widespread chronic inflammation, which is initiated by antibodies that activate the complement system.

The complement system is the primary noncellular component of the body's innate immune system, which is the nonspecific branch of the immune system.[25] The immune system is generally divided into the adaptive and innate immune systems, which are defined by their

abilities to either respond to specific pathogens or respond nonspecifically to invaders.[25] However, breaking up the immune system into these categories ignores the interconnectivity of the immune system and the ways in which the adaptive and innate systems work together to achieve the common goal of defending the body against attack.[25] It is at one of these junctions that the complement system combines the functions of the adaptive and innate immune systems to wreak havoc on the body in the disease process of SLE. The complement system is comprised of over 30 proteins which circulate in the blood and respond in specific ways to chemical signals.[25] There are three different pathways within which the complement system operates: the classical pathway, the alternative pathway, and the lectin pathway (Figure 8).[25] Each pathway is

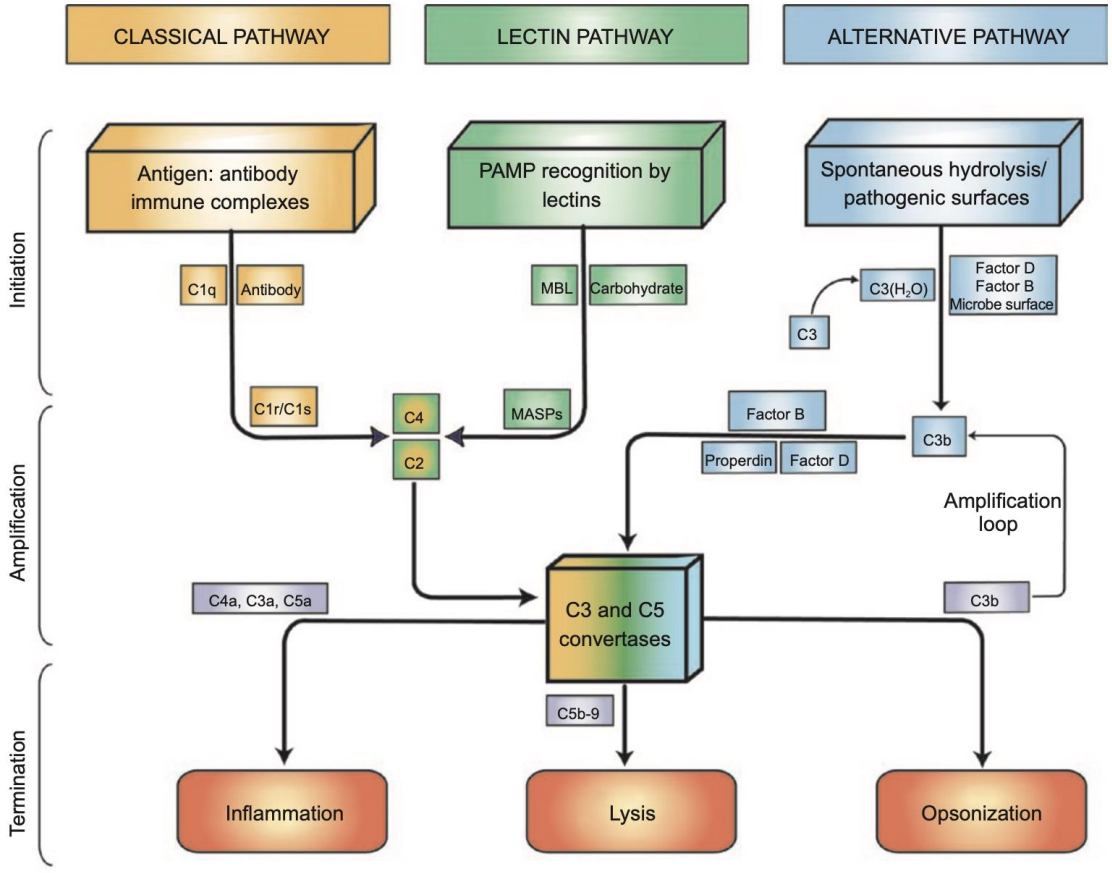

Figure 8. The complement pathways.[25]

activated by different chemical signals and each results in a different cascade reactions, but all pathways converge into a set of responses that protect the body by way of inflammation, cell lysis, or osponization.[25] It is the classical pathway that concerns SLE, because the it is activated by the binding of the first protein in the classical pathway, C1q, to the Fc region of an antibody.[25] Following the binding of C1q to the Fc region, the C1 complex forms via the addition of C1r and C1s serine proteases to C1q.[25] This complex then cleaves C4 into a small fragment, C4a, and a large fragment, C4b, along with C2 into a large fragment, C2a, and a small fragment, C2b.[25] The combination of C4b and C2a on the surface of the antigen results in the formation of a new complex called C3 convertase, which then has the ability cleave C3 into C3a, which is an anaphylatoxin, and C3b, which is an opsonin.[25] After C3b is cleaved, it then combines with the C3 convertase to form C5 convertase, which then generates C5a and C5b, which both act as anaphylatoxins.[25] The formation of C3 and C5 convertase is the point at which all three pathways converge.[25] These pathways produce three main responses (MAC assembly, anaphylatoxin/inflammatory response, and opsonization), of which the inflammatory response is central to the development of SLE.[25]

The result of the immune response to the autoantibodies generated by individuals with SLE is widespread chronic inflammation.[26] The disease pathology generated by chronic inflammation results from the effects of the body being in a long-term state intended to be an acute response.[26] The response generated by proinflammatory signals includes the expansion of blood vessels (vasodilation), increase in blood flow, capillary permeability, and migration of neutrophils into the infected tissue through the capillary wall (diapedesis).[26] Typically in acute inflammatory responses, the primary white blood cells recruited via chemotaxis are neutrophils, however during chronic inflammation, the composition of recruited cells changes to mainly

macrophages and lymphocytes because neutrophils have a short life cycle.[26] The symptoms often associated with this response are redness and swelling, since the recruitment of white blood cells requires blood vessels to become more permeable, which results in fluid accumulation at the site of the antigen.[26] The long term recruitment of macrophages, lymphocytes, and plasma cells results in tissue damage due to the mechanisms these cells use to attack antigens, and the damaged tissue is then repaired using fibrous connective tissue which fails to perform the function of the original tissue.[26] The presence of swelling causes pain in affected areas due to the pressure placed on surrounding tissues, and the replacement of the original tissues with fibrous tissues results in organs failing to preform efficiently or at all.[26] This process of excessive prolonged inflammation broadly explains the variety of symptoms experienced by patients with SLE.

Genetic & Environmental Factors

While the exact cause of SLE remains unknown, there are a number of contributing factors that have been linked to the pathogenesis of SLE.[2] There appears to be no single cause for the development of SLE, but rather there appears to be a combination of genetic and environmental factors.[2] If the cause were strictly genetic, the expectation would be that in cases where monozygotic twins carry the genetic traits to develop lupus that both would develop lupus.[27] However, this is not the case; both twins develop SLE only 24-35% of the time and the rest result in only one of the monozygotic twins developing SLE.[27] That being said, there is certainly a genetic component to the development of SLE, which can also be demonstrated by monozygotic twins.[28] Studies have shown that monozygotic twins are ten times more likely to both develop SLE than dizygotic twins, and that first-degree relatives are twenty times more likely to develop SLE than the general population.[28] This indicates that in addition to a genetic

component that predisposes individuals to the development of SLE, there must also be an environmental component that triggers the pathogenesis.

Significant amounts of research have been done attempting to discover genetic factors that might predispose someone to SLE, including many genome wide association studies (GWAS).[28] There have been over 40 different genes that have been linked to the formation of SLE with many being complex and polygenic and some being monogenic in nature.[28] One major region of the genome that is believed to play a role in disease pathogenesis is the classical human leukocyte antigen (HLA) complex, which is responsible for a wide range of genes that function in the immune system.[28] This region serves as the encoding region for the major histocompatibility complex (MHC), which enables the immune system to recognize self from non-self.[28] Some of the most prime examples of a polygenic source of disease development are the genes that regulate interferons (INFs), interferon regulating factors (IRFs).[28] Many studies have connected the dysregulation of interferons to the development of SLE, with some studies even showing that SLE can be induced by increasing levels of INF-α.[28] Other polygenic sources include genes that code for STAT4, IFIH1 and osteopontin (OPN).[28] Conversely, there have also been monogenic causes linked to SLE such as complement component C1q deficiency, three-prime repair exonuclease 1 (TREX1) or deoxyribonuclease 1-like 3 (DNASE1L3).[29] These examples only scratch the surface of the wide range of genes that have been linked to the development of SLE, and genes alone do not tell the whole story, as a genetic predisposition does not guarantee that an individual will develop SLE.

In addition to a genetic predisposition, certain environmental factors must also contribute to the development of SLE. It is these environmental factors that are the most difficult to study in connection to the development of SLE. While some environmental factors, such as ultraviolet

radiation (UVR) exposure, can be easily linked to flare-ups and disease progression, it is very difficult to connect any specific environmental factors to the onset of the disease.[30] While there is a very strong connection between UVR exposure and the development of rashes, which are among the hallmark indicators of lupus, it is difficult to say if it is the UVR exposure led to the onset of the disease or merely aggravated it.[30] Some studies haves suggested links to various other environmental factors including silica, current cigarette smoking, oral contraceptives, postmenopausal hormone therapy, air pollution, solvents, pesticides and heavy metals.[8] Environmental factors such as these have been linked to many other diseases, namely cancer, and all prove very difficult to study in relation to the beginning of the disease.[8] Among the many environmental factors linked to the development of SLE are viruses, which have been implicated in other disease processes, having even been definitively tied to the development of certain cancers.[31] For the purpose of this review viruses will be the primary environmental factor that is investigated in connection with the onset of SLE.

Viruses and Autoimmunity

The pathogenesis of systemic lupus erythematosus (SLE) has proven incredibly difficult to elucidate, and upon examination of the pathophysiology of the disease, there appears to be numerous factors that contribute to the onset of SLE. Only on rare occasions such as a C1q deficiency, does there appear to be a single aberration that leads to the development of the disease.[29] It appears that the combination of genetic factors and environmental factors leads to the development of SLE, and the particular environmental factor that appears to have the most evidence to connect it to the onset of SLE is viruses.[32] This hypothesis is based on the known mechanisms involved in viral infection and the immune response to a viral infection. This knowledge has led to the development of potential mechanisms of viral activation of autoimmunity, some of which have been supported in murine models.[33] When the mechanisms are viewed in conjunction with the known factors that contribute to the onset of SLE, the case for the viral onset of SLE becomes very compelling; with the last piece of the puzzle being the ability to test this hypothesis, which will be discussed in the following chapter.

Viruses

A brief introduction to viruses is necessary before detailing how the immune system attacks viruses. A virus is an obligatory parasite, which means that it requires a host to survive.[34] While viruses exhibit many of the characteristics of life, viruses do not qualify as living organisms.[34] Often times words used to describe viruses would seem to imply that they are alive, however they lack the ability to reproduce on their own, which is one of the criteria for life.[34] Viruses are small packages of information in the form of DNA or RNA that are surrounded by a protein coat or capsid, sometimes containing other molecular machinery such as enzymes.[34] By themselves, viruses are not capable of reproduction, but they can hijack the machinery of other

living organism to replicate themselves.[34] In this way, viruses appear to exhibit all of the characteristics of life, however they fall short only because they lack the ability to replicate without a host.

Viruses are highly specific to the host that they infect.[34] Viruses infect a host cell by binding to receptors on the surface of a cell and inserting their genetic material into the host cell.[34] This can be accomplished by either fusing the viral capsid with the host membrane or directly inserting the genetic material into the host cell while leaving the capsid outside.[34] Since the virus can only bind to specific proteins on the cell surface, typically a virus can only infect a small range of cells within a given species.[34] Occasionally a mutation in the genetic information contained within the virus can cause a significant enough modification that a virus can jump species, and these mutations typically lead to some of the more deadly viral infections that humans experience such as the human immunodeficiency virus (HIV), which was originally found in apes and monkeys.[34] The survival of a virus is dependent on its ability to infect a host and subsequently be transmitted to a new host.[34] If a virus is too efficient at infecting a host and taking over the host, the host will not live long enough to transmit the virus, and if the virus is not effective enough at infecting the host it will fail to be passed on.[34]

The "life" cycle of a virus consists of entering a host cell and integrating it's genetic material with the that of the host.[34] The virus will then enter one of two cycles: lytic or lysogenic (Figrue 9).[34] The lytic cycle consists

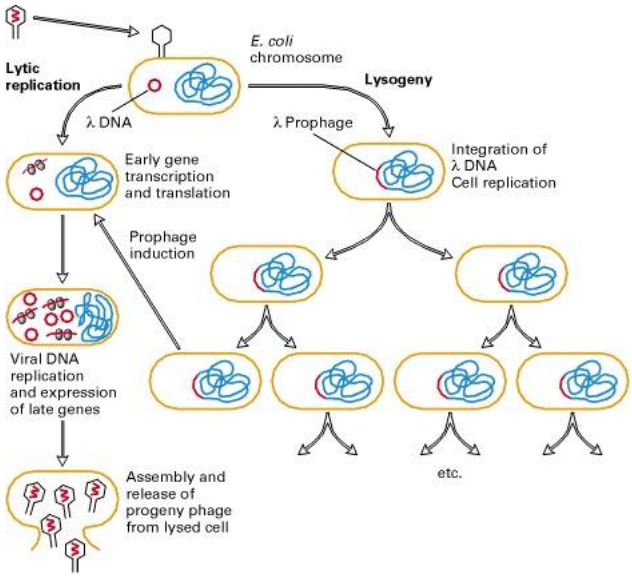

Figure 9. Viral lytic and lysogenic cycles.[34]

of allowing the host cell to duplicate the virus using its cellular machinery until there are so many viral particles in the cell that the cell bursts or lysis, allowing the replicated viruses to infect other cells.[34] The lysogenic cycle allows the cell to continue replicating itself with the viral genetic material integrated into the host's genetic material.[34] This perpetuates the viral genetic material in subsequent generations of the cell, and eventually the virus will reenter the lytic cycle and complete the steps of the lytic cycle.[34] It is the lytic cycle that is deadly for the host, since the completion of this cycle results in the death of the cell, and the viral replication increases at an exponential rate.[34] Unless the host can stop the spread of the virus, the virus will eventually kill the host by destroying entire cell populations.[34]

Human Immune Response to Viruses

The human body is equipped to handle the invasion of viruses with the immune system's various cell types and methods of protecting the host from invasion.[33] The innate immune system contains many elements that respond to viral particles, slowing down the rate of infection, and the adaptive immune system is ultimately able to develop a unique response to specific viruses that in many cases completely eliminate the virus from the body.[33] The innate immune system, which includes the complement system described previously, is the body's first line of defense against viral invasion.[33] It is not capable of responding to specific viruses, but contains elements that can react with various nonspecific elements pertaining to viruses.[33] Some of the primary cells that function as a part of the innate branch of the immune system are natural killer (NK) cells, natural killer T (NKT) cells, neutrophils, macrophages and dendritic cells (DC).[33] Macrophages and DC are among a class of cells called antigen presenting cells (APC), which can engulf viral particles and present them to cells in the adaptive branch of the immune system.[33] These cells along with NK cells have receptors on their cell surface called Toll-like receptors

(TLR), which can bind to nonspecific elements of viruses such as dsDNA, ssRNA and dsRNA.[33] The TLR initiate the release of molecules called interferons (IFN), which activate NK cells and induce an antiviral state in cells which inhibits protein synthesis rendering the virus incapable of replicating.[33] Activated NK cells can induce cell death in infected cells and release inflammatory cytokines.[33] NK cells can recognize an infected cell by the amount of MHC expressed on a cell surface.[33] Infected cells express lower amounts of MHC on their cell surface, enabling NK cells to recognize infected cells without needing to be able to respond to the specific virus.[33] In addition to the cells that contribute to the innate response, molecules such as the complement system described previously, cytokines, TNF-α, IFN-γ, IL-12, IL-6, and chemokines such as MIP-1α play a role in the innate response.[33] Despite the best efforts of the innate immune system, viruses are only slowed down by the innate branch, and it is the adaptive immune system that possesses the ability to respond specifically to viruses and completely eliminate them.

Central to the adaptive immune system's ability to recognize specific viruses are the specific receptors present on the surface of cells that function within the adaptive immune system.[35] The adaptive immune system is composed of two main cell types: B lymphocytes (B cells) and T lymphocytes (T cells).[33] B cells and T cells express highly specific receptors on their cells surface called B cell receptors (BCR) and T cell receptors (TCR) respectively (Figure 10).[35] Each immature B cell or T cell has one unique type of BCR or TCR expressed on its cell surface and throughout the body there are as many as 10^6-10^7 possible receptor variations.[35] Each of these receptors has the

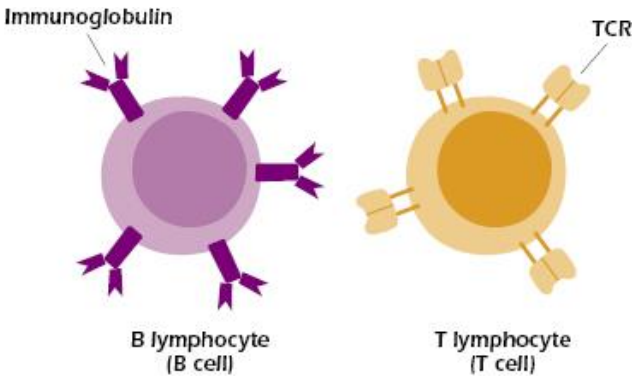

Figure 10. B cell with BCR or immunoglobulin (Ig), and T cell with TCR.[35]

potential to respond to a specific antigen, and upon encountering the right antigen with proper activation, the activated cell will clonally reproduce itself and proliferate throughout the body.[35] Equally important in the recognition of viral particles are the MHC receptors.[36,37] There are two classes of MHC receptors: MHC I and MHC II.[35] MHC receptors combine segments of the virus with the receptor to be presented on the surface of the cell.[36] Infected cells present MHC I molecules on their surface, and APC are capable of presenting antigens on either class I or II MHC molecules.[36] The MHC then corresponds with the BCR or TCR of the specific B cell or T cell that is primed to respond to a specific virus or pathogen.[36,37] These highly specific receptors allow for APC to activate an appropriate T cell response, and for the subsequent T cell response to attack appropriate cells and activate additional help.[33]

T cells are primarily responsible for eradicating viruses from the body by directly attacking infected cells and activating B cells.[36] T cells are categorized based on the receptors expressed on their cell surface and their corresponding function.[36] The two major classes of T cells are CD4 and CD8 T cells, which both express different MHC receptors on their cells surface, with CD4 T cells expressing MHC class II and CD8 T cells expressing MHC class I.[36] The MHC class dictates the immune cells with which the T cells can interact.[36] B cells only express MHC II and can only interact with CD4 T cells, while APC contain both MHC I and II enabling them to interact with both CD4 and CD8 T cells.[33] CD8 cells are also called cytotoxic T (Tc) cells, which eludes to their role in killing infected cells.[36] Upon activation by a APC, Tc cells proliferate and the TCR on the Tc cell binds to the MHC 1 receptor on infected cells.[36] Upon binding with an infected cell, Tc cells release cytotoxic granules that induce apoptosis or cytokines that have an antiviral effect and can rid the cell of viral particles without killing it.[33] CD4 T cells can be further broken up into T helper cells (Th) and T regulatory (Treg) cells, with

the Th cells playing the largest role in viral defense.[36] Th cells are necessary for the activation of B cells.[37]

The role of B cells has been previously eluded to in the discussion of antibodies and their role in the disease pathogenesis of lupus. Specifically, in the case of viral infection, B cells produce antibodies that bind to specific regions on the virus called epitopes.[38] The binding of antibodies to the viral particles both inhibits them from entering the cell and marks them for phagocytosis.[38] The activation of B cells to produce these highly specific antibodies occurs after the B cell has both encountered the pathogen and been activated by a Th cell.[37] As mentioned previously, B cells present a specific BCR on their surface which is in fact an IgM or an IgD molecule which were previously described in the discussion of antibodies.[35] Each immature B cell presents as many as 100,000 of the same Ig on its surface.[35] When a virus binds to an immature B cell BCR, the B cell takes in the virus and integrates parts of it to the MHC, and is then activated when the TCR of a Th cell binds to the MHC.[37] After the B cell is activated, a series of changes allow the B cell to class switch to produce IgG molecules, and the B cell further differentiates into either memory cells or plasma cells.[37] Plasma cells are responsible for secreting soluble antibodies into the bloodstream and these cells are proliferated throughout the body following B cell activation and differentiation.[37] The memory cells continue to survive even after the virus has been eradicated allowing the body to retain the ability to mount a defense if ever presented with the same virus.[37] Additionally some of the plasma cells retreat to the bone marrow where they continue to secrete low levels of antibodies.[37] The memory of the adaptive immune system is not limited to the B cells, with some T cells being maintained by IL-7 and IL-15 which regulate survival and proliferation respectively.[33]

It is the normal response to viruses that informs the hypothetical connection between viruses and the onset of autoimmune disease. While the immune system is typically incredibly effective at attacking and eliminating viruses, the mechanisms used to do this must be tightly regulated.[33] Furthermore, while the immune system is capable of completely eliminating most viruses, there are some viruses that can avoid complete eradication which can lead to the perpetual activation of the immune system.[33] So while in most cases the immune system does its job and does it well, there are circumstances that open the door to malfunctions which have the potential to trigger an already genetically susceptible individual.

Molecular Mimicry

There are few main mechanisms that have been proposed to explain how a viral infection can lead to the development of an autoimmune disease.[32] One of which is called molecular mimicry, which describes a molecular form of survival tactic that has been applied at larger scales throughout all forms of life.[39] Various organism have evolved to share similarities with their environment that allow them to evade detection from other organism that might seek to destroy them.[39] Infections microorganism are no exception, with many exhibiting characteristics similar to the host and allowing these microorganisms to evade detection by the host's immune system.[39] In order for a virus or another pathogen to be an effective infectious agent, it must be able to at some level avoid the host immune response; otherwise it would quickly be detected and eradicated.[39] The term "molecular mimicry" was first coined by Damian in 1964 to describe the similarities between surface markers on antigens and host cells that allowed the antigen to avoid detection.[40] Two years later, Zabriskie and Freimer demonstrated that membrane structures on group A streptococcus shared commonalities with membrane structures found on human muscle tissue.[40,41]

The proposed mechanism by which autoimmune diseases develop via molecular mimicry involves the activation of autoreactive immune cells by the pathogen bearing resemblance to the host cells.[32] In a healthy immune system, T cells are rigorously screened for self-tolerance in the thymus, and any self-reactive cells are terminated.[40] However, it is possible that these self-reactive T cells may avoid detection in a few different ways.[40] The simplest explanation is that a self-reactive T cell may simply have a TCR that is reactive to both an antigen and self, and simply fail to be detected during the screening process.[40] This type of failure may be due to genetic abnormalities.[40] Another way an autoreactive T cell could avoid detection in the thymus is the presence of two TCRs.[40] As many as 30% of T cells in the body have two TCRs with the potential for one of the TCRs to be self-reactive yet still avoid detection.[40] Lastly, it is possible that specific combinations the α and β chains may lead to a "chimera" combination that is capable of recognizing self, but somehow avoids detection in the thymus.[40] The presence of these autoreactive T cells is the basis of molecular mimicry leading to the development of autoimmune diseases such as SLE.[40] If these autoreactive T cells are present in the body, they may remain inactivated throughout the entirety of a person's life; however, if exposed to a pathogen that bears an epitope similar to the self-antigen, this can lead to the activation and proliferation of a population of T cells that are autoreactive and later the activation of B cells that release autoantibodies.[40]

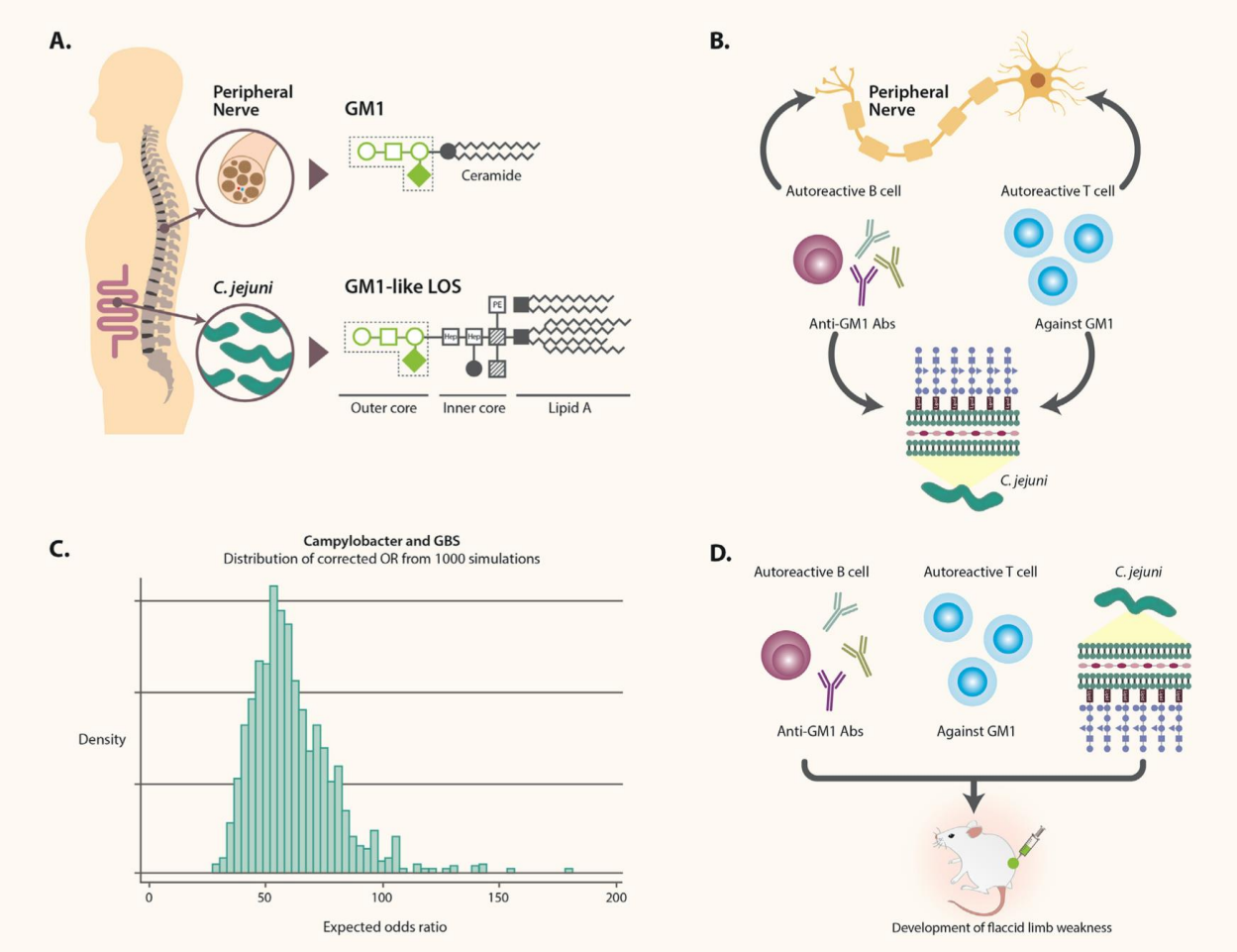

Figure 11. Criteria for identifying molecular mimicry.[40]

Three broad categories have been developed to describe the types of molecular mimicry that can occur.[39] The first category describes peptide chains that are identical in both the host and the pathogen.[39] The second type of molecular mimicry is when the epitope on the antigen is structurally similar to an epitope on a host cell, allowing for cross reactivity by the host immune system.[39] Finally, there are occasions where the completely dissimilar structures can be recognized by the same antibody, causing a rare type of cross reactivity that is not caused by similarity in structures.[39] These types of molecular mimicry have been observed in various experiments leading to the development criteria for identifying molecular mimicry.[40] There are four major criteria (Figure 11) with the first being that host and microorganism or environmental

agent share a similar epitope.[40] The second criteria require that patients with autoimmune diseases have detectable antibodies or T-cells that cross react with both epitopes.[40] For the third criteria an epidemiological link must be established between contact with the microorganism or environmental agent.[40] Finally, the fourth criteria mandates that the autoimmunity must be reproducible in an animal models exposing the subjects to the same microorganism or environmental agent possessing the epitope in question.[40] These criteria have proven difficult to demonstrate in human models, and these shortcomings have led to doubt that molecular mimicry is responsible for the onset of autoimmune diseases in humans.[40]

Despite the lack of evidence in human models to definitively link molecular mimicry to the development of autoimmune diseases, there have been several examples of animal models that have demonstrated this connection.[40] In the early 1980's work done by Fujinami et al. demonstrated that mice developed antibodies to measles virus and herpes simplex virus (HSV) that cross reacted with human cells, and later work by this group demonstrated that myelin basic protein (MBP) shared homology with the hepatitis B virus polymerase (HBVP) and after exposing rabbits to these moieties, they developed encephalomyelitis.[40] In addition to work done with animal models demonstrating the link between molecular mimicry and autoimmune diseases, numerous studies have been done providing evidence for a potential link between viruses and the development of autoimmune diseases in humans via molecular mimicry. Some examples include T cells showing reactivity to both MBP and EBV nuclear antigen 1 (EBVNA1) in patients with MS, similarities between epitopes of pancreatic β cells and viral components along with a correlation in the rise of cases of type I diabetes and enterovirus epidemics, and cross reactivity between EBVNA1 and Sm proteins in patients with SLE.[40] The difficulty with each of these examples is that none of them provide a definitive connection between the virus

and the autoimmune disease. Each provides circumstantial evidence and a plausible mechanism in molecular mimicry, but there are so many factors that contribute to the development of autoimmune diseases that current methods fall short of linking the two.

Bystander Activation

Another potential mechanism linking the development of autoimmune diseases to viral infection is called bystander activation.[32] Among the three potential mechanisms linking AD development to viruses, this mechanism has the least amount of evidence both supporting it and potentially linking it to the development of SLE.[42] In fact, evidence suggest that this mechanism could potentially have a protective effect in SLE by activating T cells with an immunosuppressive effect.[42] Despite the lack of evidence, there still remains theoretical potential due to the possible activation of autoreactive T cells via this mechanism. Each mechanism is ultimately reliant on the activation of autoreactive T cells that are already present in a susceptible individual, with each describing a mechanism by which a viral infection could lead to this activation.

The mechanism by which bystander activation works can be described in two different ways, with one lending similarity to epitope spreading, which is the final mechanism that will be discussed. The primary definition of bystander activation is the activation of bystander T cells that does not involve the TCR.[42] This mechanism proposes that during a viral infection either memory T cells or naïve T cells can be activated by soluble mediators released by T cells actively fighting the infection.[42] It has been demonstrated that CD8 T cells can be activated by soluble mediators such as IL-12, IL-15, IL-18, and type I INFs.[43] It is proposed that during a viral infection, the T cells fighting the infection can activate T cells which are not specific to virus by bystander activation through the release of these soluble mediators.[42] These T cells

activated through bystander activation have been shown to contribute to inflammation though the production of inflammatory cytokines such as INF-γ.[43] Additionally, this mechanism has been shown to work similarly in memory-like CD4 T cells in response to the release of IL-1β and IL-23.[44] In a study on the role of bystander activation in autoimmune encephalomyelitis these CD4 T cells were shown to contribute to inflammation via the production of IL-17A and IFN-γ.[44] The alternate way in which bystander activation is described overlaps with the description of epitope spreading, which needs to be described separately.

Epitope Spreading

Out of the three mechanisms, epitope spreading aligns the closest to the pathogenesis of SLE. The known factors that contribute to the pathogenesis of SLE are impaired apoptotic clearance and self-reactive lymphocytes which lead to the development of autoantibodies and widespread inflammation.[29] The basis of epitope spreading is that a prolonged viral invasion leads to the exposure of cellular components that activate autoreactive lymphocytes, and subsequently lead to the spreading of lymphocytes that are reactive to self-epitopes.[45] There are two routes that can lead to the recognition of self-epitopes by lymphocytes during a viral infection, and both are initiated by the mechanisms that immune system normally uses to fight a viral infection.[45]

The first route of epitope spreading begins with the influx of lymphocytes to an infected area, which in a normal immune response involves Tc cells inducing apoptosis in infected cells.[45] The influx of lymphocytes in addition to the pervasive cell death can lead to the introduction of autoreactive lymphocytes to intracellular components.[45] Since the normal immune response leads to the recruitment of lymphocytes, this increases the likelihood that an autoreactive lymphocyte will be recruited to an area of viral infiltration. The method by which Tc cells eradicate the virus

can then lead to the exposure of epitopes that would not normally be found in such large quantities.[45] These two factors can lead to the right sequence of events for autoreactive lymphocytes to proliferate following the exposure to epitopes that were present in the wake of an immune response to a viral infection.[45]

The second route by which epitope spreading can lead to the development of autoimmunity is mediated by APCs (Figure 12).[45] This route closely resembles bystander activation, and some descriptions of bystander activation will match this description of epitope spreading.[32,45,46] The main difference between the typical definition of epitope spreading and the similar definition of bystander activation is epitope spreading is generally connected to a chronic viral infection, and bystander activation is generally associated with an intense acute viral infection.[32,45,46] The mechanism ends up being incredibly similar and the name becomes inconsequential. The difference between the second route of epitope spreading and the first is the presentation of the epitopes by APCs as opposed to the lymphocytes being directly exposed to the epitope.[45] In this route, the persistent availability of intracellular components leads to these components being taken up by APCs and then the epitopes of these components are integrated into the MHC of the APC.[45] The APC then presents these epitopes to self-reactive lymphocytes leading to the proliferation of these lymphocytes and the eventual production of autoantibodies, which is the major contributor to the pathophysiology of SLE.[20,45] This particular route of epitope spreading

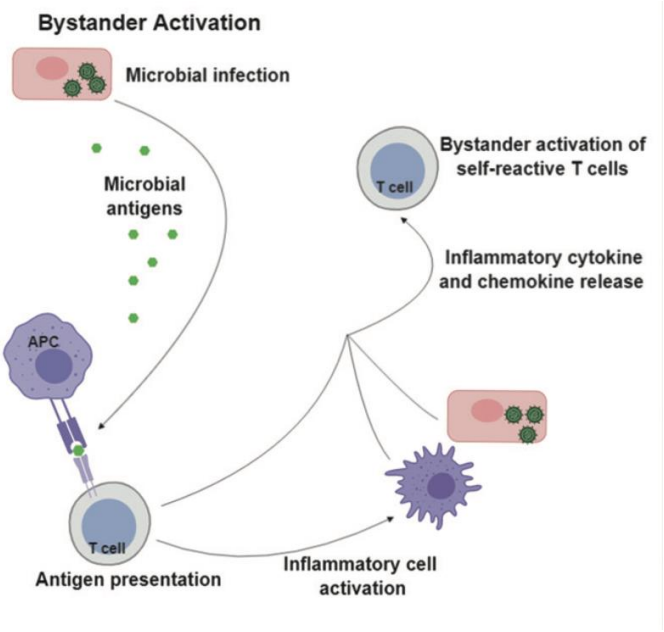

Figure 12. Bystander activation via APC.[45]

has the potential to be even more harmful than the first. The uptake of self-molecules by APCs and subsequent presentation to lymphocytes can lead to the recognition of multiple self-epitopes as a result of the process by which epitopes are integrated with MHC.[45] When APCs engulf a self-molecule, they can integrate more than one epitope from that molecule to the MHC.[45] This creates the potential for epitopes that may be shared by more than one self-molecule to be recognized, leading to the recognition of multiple self-molecules by autoreactive lymphocytes.[45] The uptake of one self-molecule by an APC can then lead to the recognition of another and potentially then leading to the production of a variety of autoantibodies, a hallmark of SLE.[45]

As with the other proposed mechanisms of viral induction of autoimmunity, the evidence supporting this mechanism is primarily in animal models. One popular example is the persistent infection of Theiler's virus in mice, which leads to the development of an experimental autoimmune disease, experimental allergic encephalomyelitis (EAE).[47] A more recent example of experimental evidence for epitope spreading is an experiment on mice that showed epitope spreading in experimental autoimmune uveoretinitis (EAU).[48] The obvious limitations with both of these examples is the experimental nature of these autoimmune disease and their restriction to mice. However, studies have been done linking multiple viruses and epitope spreading in humans to SLE, which provides evidence that this mechanism could explain the pathogenesis of SLE.[49–51]

Connecting Viruses to Autoimmune Diseases

Each of these three mechanisms provide plausibility to the hypothesis that viral infections could play a pivotal role in the development of an autoimmune disease such as SLE. The difficulty in proving this connection has always been the limitations in experimental design. In order to provide more conclusive evidence regarding this connection a study would have to

figure out both the identity of the virus causing the autoimmune disease and what autoimmune disease(s) that virus causes. A study of this size and scope has not previously been possible with available technologies as there are over 200 different viruses that commonly infect humans and there are almost 100 different autoimmune disease.[52,53] Even focusing on just one autoimmune immune disease requires selecting only one or a few viruses to investigate, which amounts to a shot in the dark. Any evidence that is acquired requires enormous amounts of additional data to eliminate all other explanations for the results obtained from such a study. A comprehensive study would require the ability to test people for every virus with which humans are commonly infected and a longitudinal analysis of the autoimmune diseases developed after viral exposure. VirScan has been recently developed and could accomplish this goal.

VirScan

In science, every technology is built on the foundation of previously discovered technologies. Observations that had been made in the past can be taken and applied to propose new hypotheses, which lead to new discoveries. Without the former, the latter would not exist. The technology used to test the hypothesis of my paper is no different. To understand the significance of this technology, it must be broken down into its essential components. After understanding how these components work, the application becomes both easier to understand and more obvious. What the technology accomplishes is simple to explain, but how it accomplishes it and the application of it are less simple. VirScan is a combination of several different techniques that makes it possible to test a small sample of a person's blood and determine what viruses that person has likely been exposed to during their lifetime.[54] The central component that makes this possible is called Phage Immunoprecipitation sequencing (PhIP-seq), which is in of itself the summation of several different techniques that accomplish a central goal.[54] The supplemental technology that provides the efficiency and financial feasibility of creating and using a technology like PhIP-seq is owed to the advances in DNA sequencing and analysis, *de novo* DNA synthesis technologies, and development of T7 bacteriophage libraries.[55]

High Throughput DNA Sequencing

The fundamental aspect of identifying viruses is the genetic code. Viruses are coded with either ribonucleic acid (RNA) or deoxyribonucleic acid (DNA), which dictate what the virus looks like, what kind of cells it can infect, and virulence. Reading this code is central to understanding the identity of a virus as well as understanding its components. The genetic code consists of five nitrogenous bases that each correspond with a letter: adenine (A), guanine (G), cytosine (C), thymine (T), and uracil (U).[56] Theses bases are connected to one another by a sugar

phosphate backbone.[56] The type of sugar that comprises the backbone in conjunction with the bases that make up the genetic code differentiate the make-up of DNA from RNA.[56] DNA uses thymine instead of uracil and RNA does the opposite, which means that the genetic code for both is comprised of four-letter units and can be translated from one to the other. When translating between the two, uracil corresponds with thymine.[56] The order in which these bases are arranged along a DNA or RNA strand determines the various characteristics of the virus. Dr. Fredrick Sanger earned a Nobel prize in Chemistry in 1980 for developing a method for reading a DNA sequence.[57] His method uses fluorescently labeled 3′-dideoxynucleotides (ddNTPs) to cause identifiable termination in DNA elongation.[57] DNA is normally replicated using DNA polymerases, which are enzymes that are able to duplicate an entire strand of DNA.[56] His method first uses DNA polymerase to amplify the DNA strand that is being read and once there are numerous copies, the DNA is allowed to copy with ddNTPs present.[57] Whenever a ddNTP is entered into the growing DNA strand, elongation is terminated.[57] This process produces DNA strands of varying lengths, from which the DNA code can be elucidated.[57] While the development of this technology was groundbreaking, it was also very limited. It cost over a billion dollars and took almost a decade to sequence a human genome using Sanger's method.[57] The Sanger method would be far too expensive and require too much time to scan libraries of viral genomes.

VirScan scans libraries of viral genomes to determine an individual's exposure to different viruses.[54] The technology that makes it possible to read and analyze the number of DNA sequences used by VirScan is called high throughput sequencing (HTS) and is capable of sequencing an entire human genome in one or two days for $1,000 or less.[57] The process is several magnitudes cheaper and shorter than Sanger's initial method, but is built on the same

principles. The leading company in HTS, Illimuna, uses a bridging technique to achieve fast and inexpensive DNA sequencing (Figure 13).[57] This technique employs oligonucleotide (oligo) strands that are anchored to glass flow cells.[57] After splicing the original DNA sequence into many smaller sequences, the smaller sequences are modified with identifiers and sequences that complement the oligo strands in the flow cells.[57] Once attached to the oligo strands, these DNA segments are replicated using a process called bridge amplification.[57] Once the strands have been replicated hundreds of thousands of times, these strands are read using fluorescently tagged nucleotide bases.[57] The order of the DNA sequence can then be reconstructed by overlapping sequences and computational analysis of the data collected during the addition of the fluorescent tags.[57] Most of this process is automated and requires very little manual labor.[57]

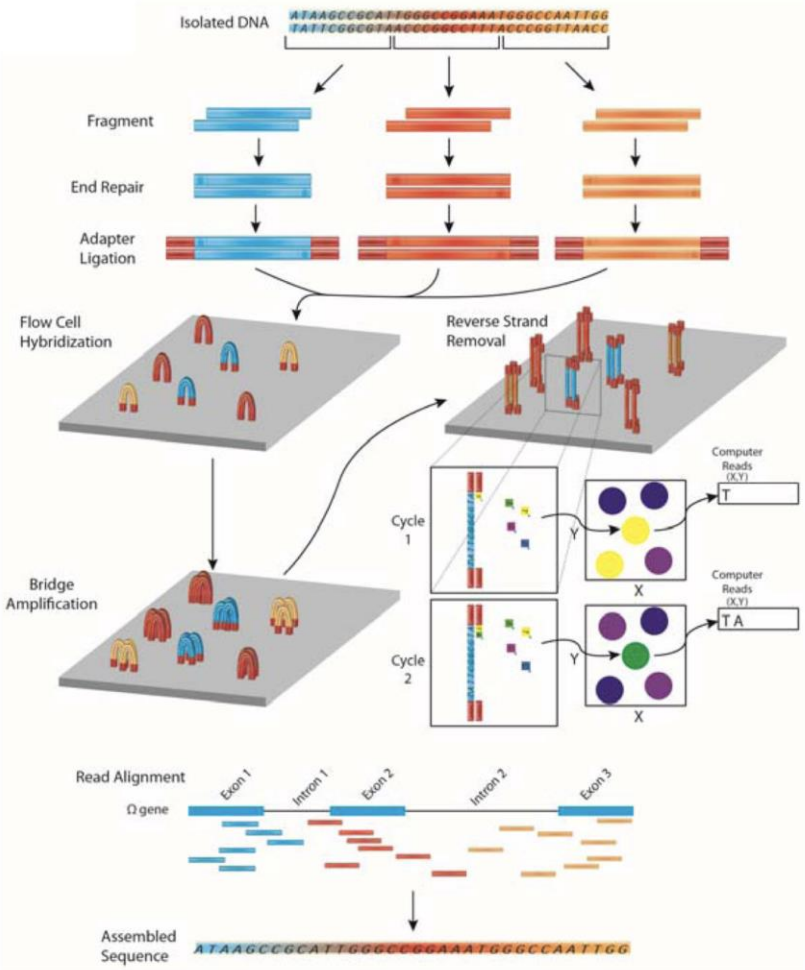

Figure 13. Overview of DNA-sequencing using the Illumina platform.[57]

De Novo DNA Synthesis

In addition to being able to read the genetic code of viruses, the VirScan technology also requires the ability to write a synthetic genetic code. The technique for synthetically creating

oligo strands was developed in the 1950s by Todd, Khorana, and their colleges.[58] At its core, it is a reaction that forms a phosphodiester bond between two nucleosides, which in isolation, is not a difficult reaction to achieve.[58] The difficulty of synthesizing a DNA strand of a desired sequence, is controlling the order in which the nucleosides are added.[58] Controlling which nucleosides are added is achieved in four steps (Figure 14).[58] The first step is the attachment of a dimethoxytrityl (DMT)-protected nucleoside phosphoramidite to a solid support and then deprotecting it using trichloroacetic acid.[58] The second step is to add the desired DMT-protected nucleoside to the newly unprotected 5′ hydroxyl group, which forms a phosphodiester bond.[58] The 5′ hydroxyl group can then be capped by acetylation to prevent degredation.[58] Finally, the phosphite backbone is oxidized via iodine oxidation resulting in a cyanoethyl protected phosphate backbone.[58] The process is started again by removing the DMT, creating a cycle that can be repeated until the desired sequence has been obtained.[58] Once this process has been completed, the oligo can be removed from the solid support, and the backbone and bases can have their protecting groups removed.[58]

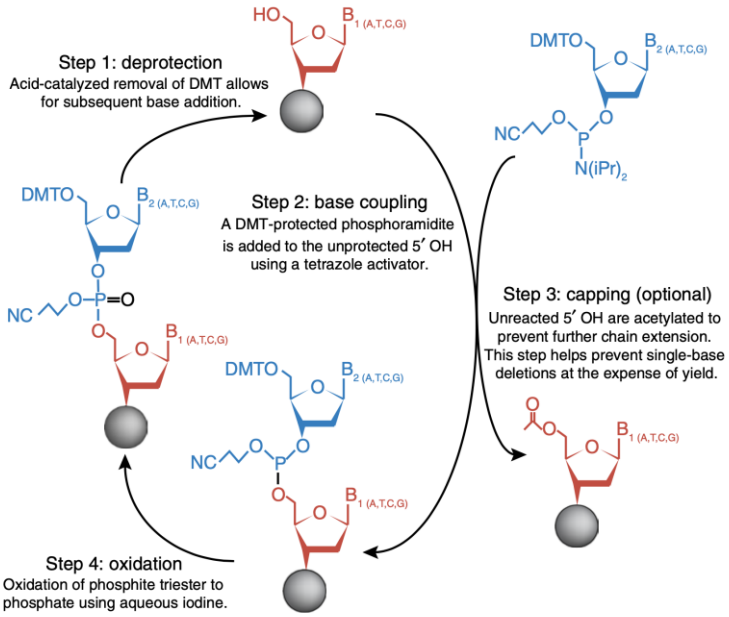

Figure 14. Four step synthetic oligo synthesis.[58]

It was not until HTS became widely available and paved the way for understanding the language of DNA, that the technology for writing original DNA became relevant for research

purposes.[58] It is no use writing in a language, if what is being written is unknown. As with initial methods for DNA sequencing, the originally developed methods for synthesizing oligos were not efficient enough or accurate enough to be applied on a large scale like what is necessary for VirScan. Current methods use ink jet microarrays that are capable of applying exact amounts of reagents to microarrays that ensure the addition of only the desired nucleoside and achieve a high level of proficiency in carrying out the desired reaction.[58] As a result of such proficiencies, entire viral genomes can be recreated synthetically such as the 1918 Spanish influenza virus.[58] This technology also opens the door to designing viral genomes to carry out specific functions.

T7 Phage Display Systems

Central to the VirScan technology is the development of bacteriophages that can be modified to display desired surface antigens. In 1985, George Smith discovered that gene III in a filamentous phage was responsible for encoding a minor coat protein, and this gene could be modified to express foreign peptide sequences on the virion capsid.[59] His work was further modified using various phages such as f1, fd, T4, M13 and T7.[60] Most modern research employs the M13 and T7 phages, with the T7 phage being the desired option for the VirScan technology.[54,60] The T7 phage system offers several advantages over the M13 system which directly apply to how it is employed for the VirScan technology.[60] M13 phage display systems take longer to develop and are more limited in the size of display that is inserted.[60] The structural differences between M13 and T7 phages allow the T7 phages to more easily express the desired peptide sequence on the surface of the phage and allow for easier insertion of the foreign cDNA.[60] Bacteriophage T7 is one of seven different viruses that can infect and replicate within *E. coli*, which makes it a particularly easy vector to use because of the relative ease of growing and maintaining *E. coli* cultures.[60] The structure of the T7 phage includes an outer shell that is

composed of proteins gp10A and 10B, a tapered internal cylinder composed of proteins gp6, 7, 14, 15 and 16, a connector composted of gp8, a tail composed of proteins gp7, 3, 11 and 12, and tail fibers composed of gp17 (Figure 15).[60] The location that the desired peptide sequences can be inserted is at the C-terminus of the gp10B capsid protein which is the secondary capsid protein and is expressed at a 1:9 ratio compared to the primary capsid protein, gp10A.[60] This allows T7 phage libraries to be customized to express numerous foreign surface proteins without disrupting the functionality of the phage.[60]

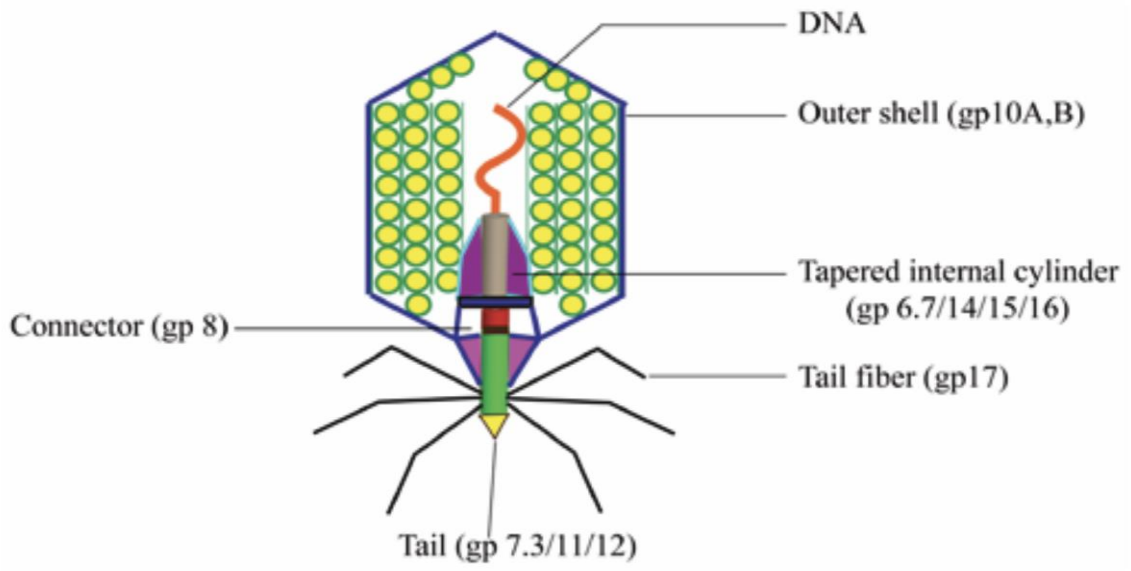

Figure 15. Detailed structure of T7 phage.[60]

Phage Immunoprecipitation Sequencing (PhIP-seq)

The PhIP-seq technology was developed by researchers at Harvard and was later slightly modified and then applied to develop the VirScan technology.[54,55] At the foundation of this technology is the biological principle of antibodies and antigens. As described previously, the body makes specific antibodies that respond to antigens on specific pathogens. It is then possible to use these antigens to identify what pathogens an individual has been and is currently exposed

to based on what antibodies are present in the body. The key to using antibodies to identify specific pathogens is identifying the specific epitope on the antigen to which the antibody binds.[61]

The PhIP-Seq technology combines the technologies of *de novo* DNA synthesis, T7 bacteriophage libraries and HTS to create a system that is capable of identifying an individual's exposure to specific pathogens using the antibodies present in the individual's blood (Figure 16).[55] The first step of PhIP-Seq is choosing a pathogen to be identified.[55] Then the epitopes to which antibodies bind have to be identified.[55] Once the epitopes have been identified, the segment of DNA that codes for the peptide sequence of the epitope must be identified.[55] After the epitopes and their corresponding DNA sequences have been identified, the DNA must be synthesized and then amplified for fusion with a T7 phage vector.[55] The T7 phage vector will then express the epitope on its surface.[55] The expression of the epitope in a vehicle other than the original pathogen is the greatest limiting factor of this technology.[55] The T7 phage vector is only capable of expressing linear epitopes, but there are several other types of epitopes which antibodies bind to that cannot be expressed on a T7 phage.[55] Examples of epitopes that cannot be expressed using a T7 phage vector are epitopes longer than 90 amino acids, epitopes containing post translational modifications, epitopes with disulfide bonds, and epitopes that are discontinuous.[55] Although this limitation means that this technology cannot currently be used to test every known epitope, linear epitopes provide enough material to make this technology useful and widely applicable.[55] Once the T7 phage library has been developed, the phage library is introduced to a sample containing antibodies which will then bind to any phage vectors expressing the epitope to which the antibody is specifically adapted. [55] Once the antibodies are allowed to bind to the phage vectors, magnetic beads coated with A/G proteins that will bind to

the antibodies are used to precipitate the bound vectors.[55] The identity of the bound vectors is then revealed using HTS.[55] The information obtained using HTS can then be used to determine the specific pathogen from the antibody sample as well as the amount of antibody present in the sample.[55]

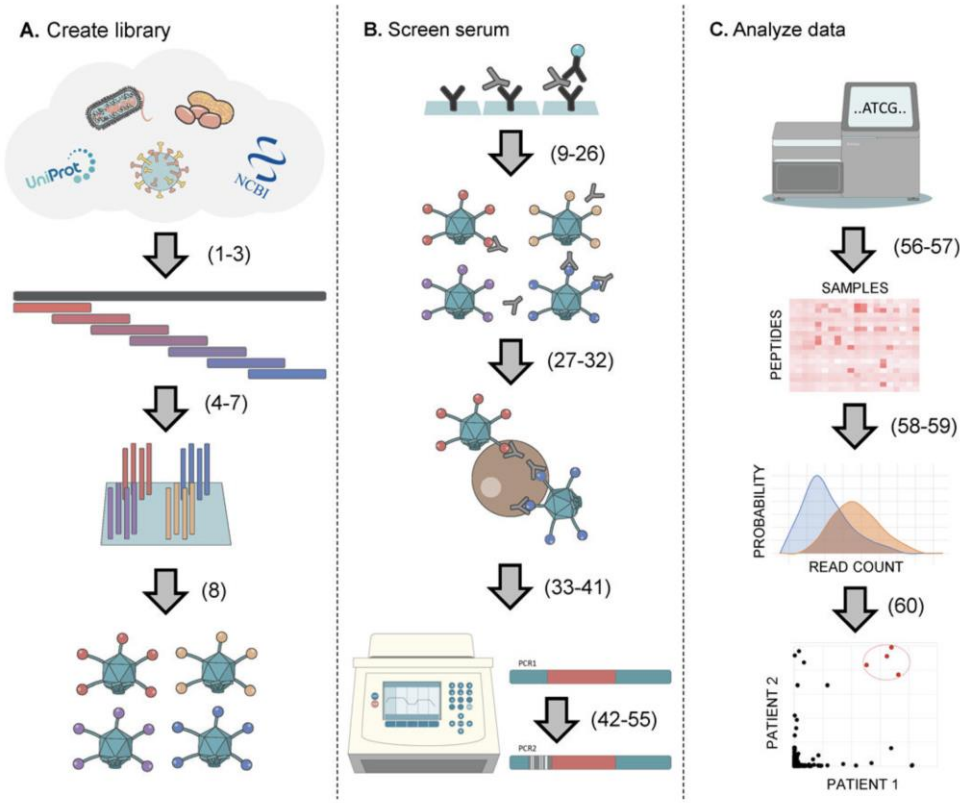

Figure 16. Overview of the PhIP-seq methodology.[55]

VirScan Design

The PhIP-Seq technology was modified to specifically identify known human viruses that a person may have been exposed in their lifetime, and this modification has been developed into VirScan.[54] In the paper outlining VirScan's proof of concept, researchers at Harvard tested 569 people from four continents, which amounted to the screening of over 10^8 antibody-peptide interactions.[54] This technology opens the door to study the effects of viruses that cause both acute and chronic infection, contributing to what is called the human virome.[62] The human

virome describes all the different viruses that exist within a person, including those that infect human cells causing disease and those infecting microorganisms that live on or in the host.[62] The current ability to test the human virome and understand the effects of a specific virus on the human body is limited to doctors making educated guesses.[63] Current technologies only allow for one virus to be tested for at a time, which is a long and expensive process that is often not worth the effort without conclusive evidence and is one of the primary limiting factors to linking a virus to the onset of SLE.[63] While the technologies for testing specific viruses are becoming more efficient, they are often times not widely available, and are still limited in their ability to enhance research into the human virome.[64] These tests are limited for research purposes because they only test for one virus at a time and often require an active infection to be successful.[64]

To construct a test capable of examining an individual's lifetime viral exposure, the T7 phage library was designed to include epitopes from all available information on viruses known to infect humans (Figure 17).[54,63] To accomplish this task, Xu et al. constructed a library of oligonucleotide sequences containing 93,904 different sequences, each being 200 nucleotides in length and coding for proteins with 56-residues and 28-residue overlaps.[54] This library was compiled using information from the UniPort database.[54] This compilation of viral data contains

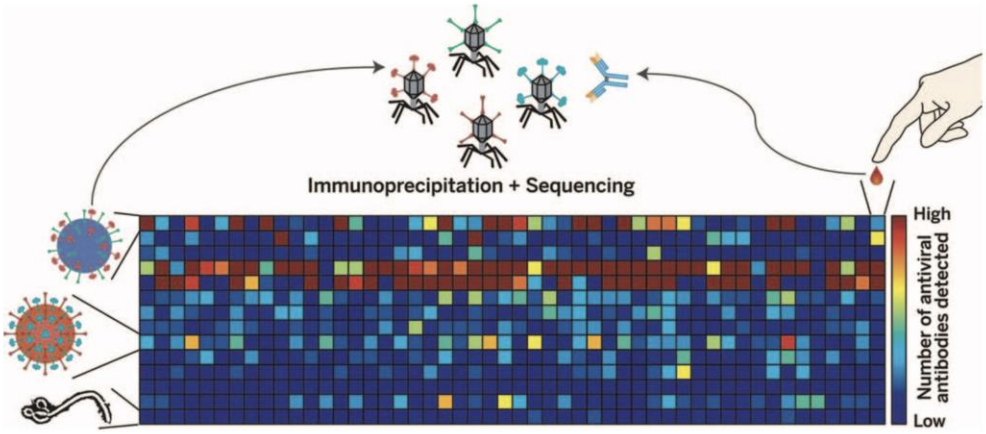

Figure 17. Systematic viral epitope scanning (VirScan).[63]

peptides from 206 species of virus and over 1,000 different strains.[54] The vast array of viral epitopes displayed on this T7 phage library provides an unparalleled amount of information pertaining to an individual's viral history.[54] The difficulty with collecting this much information is analyzing it and interpreting it in a way that translates to useful information.

The construction of the T7 phage library is the main modification to the PhIP-Seq technology that differentiates the two. The rest of the procedure follows the PhIP-Seq protocol. After the phage library is created, the serological sample is added to the display library. The phage particles that attach to immunoglobulins from the sample are precipitated using magnetic A/G beads, and HTS is used to analyze the reactive phages. The interpretation of the data obtained is the other distinguishing factor that separates the VirScan technology from the standard PhIP-Seq protocol.

The information obtained from a VirScan test must be interpreted after the procedure has been completed. A large portion of the work done by the research team at Harvard to develop the VirScan test into an effective diagnostic tool was taking the data obtained from the protocol and giving the data meaning.[54] Xu et al. developed a computational method that translates the data obtained from HTS into a test that connects the reactive phage particles with a specific virus and determines if the number of reactive phage particles corresponds with a positive or negative test for that virus.[54] Existing data on population exposure to different viruses was used to assess the accuracy of the test. Such data can be used to correct and refine the test so that the accuracy and specificity of the test can be increased.[54] For example, 29% of the samples tested positive for Cowpox virus, but upon further examination, the peptide sequence of the Cowpox virus that was reactive with antibodies from the samples is highly similar to a peptide sequence found on the Clumping Factor B protein from *Staphylococcus aureus*.[54] Finding peptide sequences that have

similar or overlapping sequences with other known antigens is essential to making the test more specific.[54] The researchers also discovered that there were certain peptide sequences that were commonly recognized by the patient population which could be used to improve the accuracy of the test.[54] For the respiratory syncytial virus (RSV), the original VirScan results were much lower than expected, which was verified by using a more accurate test that uses antibodies to detect a single virus called an enzyme-linked immunosorbent assay (ELISA).[54] This test showed that 95% of the patients tested positive for this virus, but VirScan only showed 63% testing positive for RSV.[54] After adapting the protocol for a positive test using information about commonly recognized sequences, the test results improved to 97% positive.[54] This shows that the VirScan test can be further modified to increase specificity.[54]

The novelty of this technology has led to limited applications thus far, with only nine published research articles having cited VirScan as a diagnostic tool used in the research at the beginning of 2020. The ways in which the VirScan technology has been employed can be grouped into three broad categories. The first way that VirScan has been used is to measure the activity of the immune system. One study used VirScan to measure the antibodies present in the blood of individuals before and after exposure to the measles virus.[65] While another study used VirScan to measure the antibodies present before and at different points after hematopoietic cell transplantation.[66] The last study that used VirScan to measure immune function, used VirScan to measure the activity of B cells following CD19-directed chimeric antigen receptor T-cell immunotherapy on B cell malignacies.[67] Another study used VirScan to monitor patients following kidney transplants, and found VirScan to be highly reliable and cost effective method of monitoring patient post-surgery.[68] This specifically shows that VirScan can be employed as a cost effective screening method for procedures that could be complicated by the presence of

certain viral infections. The final application of VirScan that has been demonstrated in research thus far is the implication of a virus in a disease process that could not previously be conclusively connected to a specific virus. It is especially difficult to link viruses to neurological disease processes due to the difficulty of testing for viral particles in the cerebrospinal fluid (CSF).[69] In two separate studies, VirScan was used to implicate the enterovirus A71 (EV-A71) as the cause of pediatric acute flaccid myelitis (AFM).[70,71] This application of VirScan is especially promising due to the difficult nature of implicating viruses in chronic disease processes. VirScan has the potential to link previously idiopathic disease processes to viral causes due to its ability to reveal patients' viral history even without a pervasive infection or high levels of circulating antibodies. Additionally, VirScan also introduces the ability to screen for a number of viruses that was previously impossible from a practical perspective. Prior to the development of VirScan, viruses had to be scanned for individually and the introduction of this technology introduces the ability to scan for over 200 viruses at once. This enables for a comprehensive study examining the entire viral exposure of individual in one cost effective test, which could make possible the investigation of a viral cause to the onset of SLE.

Conclusion

A comprehensive analysis of the evidence regarding SLE pathogenesis points to a genetic predisposition to autoimmunity that is potentially triggered by environmental factors. In the same way that a Rube Goldberg machine has all the pieces in place to set off a complicated series of events, individuals with a genetic predisposition to SLE appear to have many genetic components that can result in autoimmunity; however, if the right event does not take place to set off the series of interconnected pieces, the disease never develops. The difficulty of connecting any particular environmental factor to the onset of SLE is twofold. The first being that the onset of SLE is slow and appears to have a preclinical asymptomatic phase, an incomplete onset phase, and finally a complete onset phase.[29] This makes connecting any environmental factors to the onset of the disease incredibly difficult, because the exposure to the environmental factor could have occurred long before the manifestation of the disease; this essentially renders it impossible to connect the two events, as there has been no way to retroactively go back and connect the dots. Secondly, designing a method of testing environmental factors for their connection to the onset of the disease has proven to be a difficult undertaking.[33] The scale and precision of an experiment needed to link the two has thus far proven to be impossible to construct, and any attempts thus far to connect the two have not provided conclusive enough evidence.[33] This sets the stage for the promise of VirScan being employed as a diagnostic tool for such a study.

Past experiments have always been limited to testing for one or a handful of viruses due to the limitations of existing technologies. This made any correlations inconclusive because other viruses could not be eliminated, and it also limited the scope of any study attempted, because it would not be worth investigating in depth due to the fact that the results would be inconclusive at the end of the study regardless. The development of VirScan eliminates that greatest limiting

factor by enabling testing for almost every virus that is known to infect humans in one cost effective test. This then presents the opportunity to increase the scope of a study investigating the connection between viruses the development of SLE, because the results obtained would provide incredibly useful and insightful data regarding the correlation between viral exposure and subsequent autoimmune disease development. In theory, a longitudinal study could definitively determine the entire viral exposure for a massive population of subjects over a period of time and follow the development of any autoimmune diseases such as SLE. If a specific virus or family of viruses prove to be the causative agent for the development of SLE, a vaccine could serve as a preventative measure that keeps people from developing SLE.

The literature currently available on the pathology and pathogenesis of SLE strongly suggest that an environmental factor such as a virus could serve as a trigger for SLE in genetically susceptible individuals. Additionally, available literature on the correlation between viral exposure and autoimmunity strongly suggests that the mechanisms exist to initiate a cascade of events that could then lead to the development an autoimmune disease such as SLE. Finally, there now exists a technology that makes possible an investigation into this theoretical connection. The literature reviewed for the purpose of this thesis seeks to provide a solid foundation on which such a study could be reasonably attempted. In conclusion, the evidence obtained and laid out throughout this paper demonstrates that the research that has been conducted thus far investigating the pathogenesis of SLE correlates strongly with research that seeks to specifically connect viruses to the pathogenesis of SLE. Finally, the literature also shows that a technology now exists that had not previously been available to test this correlation, which could then be applied to test the hypothesis of a virus being the causative agent in the onset of SLE.

With any disease that is idiopathic in nature, it is critical to determine what leads to the development of the disease and if it can be prevented. Otherwise, as with SLE, the only solution is to treat symptoms. When diseases are caused by genetics, current technologies fall short of cures in most cases. However, when caused by an environmental agent, disease prevalence can more easily be reduced or even prevent. With SLE, evidence strongly points towards an environmental cause, and available literature provides strong evidence for a correlation between that the disease pathology of SLE and mechanisms of interaction between viruses and the immune system. The previous limitations of research investigating this connection can be overcome using VirScan to test large populations of people for their exposure to viruses known to infect humans. This then provides the means to supplement available data that points to viruses being a causative agent in the development of SLE. Ultimately, this then provides a direct line to answering the question presented in this thesis, “Can viruses cause lupus?” and answering this question leads could lead to the prevention of SLE.

References

1. Impact on Daily Life | Lupus Foundation of America. Accessed October 3, 2020. http://www.lupus.org/understanding-lupus/impact-on-daily-life

2. Maidhof W, Hilas O. Lupus: An Overview of the Disease And Management Options. *Pharm Ther*. 2012;37(4):240-249.

3. Systemic Lupus Erythematosus (SLE) | CDC. Published October 18, 2018. Accessed October 3, 2020. https://www.cdc.gov/lupus/facts/detailed.html

4. Lupus facts and statistics | Lupus Foundation of America. Accessed October 3, 2020. http://www.lupus.org/resources/lupus-facts-and-statistics

5. Carter EE, Barr SG, Clarke AE. The global burden of SLE: prevalence, health disparities and socioeconomic impact. *Nat Rev Rheumatol*. 2016;12(10):605-620. doi:10.1038/nrrheum.2016.137

6. Drenkard C, Lim SS. Update on lupus epidemiology: advancing health disparities research through the study of minority populations. *Curr Opin Rheumatol*. 2019;31(6):689-696. doi:10.1097/BOR.0000000000000646

7. Wallace DJ, Hahn B, Dubois EL, eds. *Dubois' Lupus Erythematosus and Related Syndromes*. 8th ed. Elsevier/Saunders; 2013.

8. Parks CG, de Souza Espindola Santos A, Barbhaiya M, Costenbader KH. Understanding the role of environmental factors in the development of Systemic Lupus Erythematosus. *Best Pract Res Clin Rheumatol*. 2017;31(3):306-320. doi:10.1016/j.berh.2017.09.005

9. Cojocaru M, Cojocaru IM, Silosi I, Vrabie CD. Manifestations of Systemic Lupus Erythematosus. *Mædica*. 2011;6(4):330-336.

10. Aringer M, Leuchten N, Johnson SR. New Criteria for Lupus. *Curr Rheumatol Rep*. 2020;22(6). doi:10.1007/s11926-020-00896-6

11. Lupus facial rash. Mayo Clinic. Accessed August 5, 2020. https://www.mayoclinic.org/diseases-conditions/lupus/multimedia/lupus-facial-rash/img-20007730

12. Discoid Lupus Erythematosus - American Osteopathic College of Dermatology (AOCD). Accessed August 5, 2020. https://www.aocd.org/page/DiscoidLupusErythe?

13. Chi AC, Neville BW, Krayer JW, Gonsalves WC. Oral Manifestations of Systemic Disease. *Am Fam Physician*. 2010;82(11):1381-1388.

14. Justiz Vaillant AA, Goyal A, Bansal P, Varacallo M. Systemic Lupus Erythematosus (SLE). In: *StatPearls*. StatPearls Publishing; 2020. Accessed August 4, 2020. http://www.ncbi.nlm.nih.gov/books/NBK535405/

15. Elkon K, Casali P. Nature and functions of autoantibodies. *Nat Clin Pract Rheumatol*. 2008;4(9):491-498. doi:10.1038/ncprheum0895

16. Fava A, Petri M. Systemic Lupus Erythematosus: Diagnosis and Clinical Management. *J Autoimmun*. 2019;96:1-13. doi:10.1016/j.jaut.2018.11.001

17. Schroeder HW, Cavacini L. Structure and Function of Immunoglobulins. *J Allergy Clin Immunol*. 2010;125(2 0 2):S41-S52. doi:10.1016/j.jaci.2009.09.046

18. Sathe A, Cusick JK. Biochemistry, Immunoglobulin M (IgM). In: *StatPearls*. StatPearls Publishing; 2020. Accessed August 4, 2020. http://www.ncbi.nlm.nih.gov/books/NBK555995/

19. Vidarsson G, Dekkers G, Rispens T. IgG Subclasses and Allotypes: From Structure to Effector Functions. *Front Immunol*. 2014;5. doi:10.3389/fimmu.2014.00520

20. Dema B, Charles N. Autoantibodies in SLE: Specificities, Isotypes and Receptors. *Antibodies*. 2016;5(1). doi:10.3390/antib5010002

21. Elmore S. Apoptosis: A Review of Programmed Cell Death. *Toxicol Pathol*. 2007;35(4):495-516. doi:10.1080/01926230701320337

22. Yang F, He Y, Zhai Z, Sun E. Programmed Cell Death Pathways in the Pathogenesis of Systemic Lupus Erythematosus. *J Immunol Res*. 2019;2019. doi:10.1155/2019/3638562

23. Apoptosis. Genome.gov. Accessed August 5, 2020. https://www.genome.gov/genetics-glossary/apoptosis

24. Pravda J. Systemic Lupus Erythematosus: Pathogenesis at the Functional Limit of Redox Homeostasis. *Oxid Med Cell Longev*. 2019;2019:1651724. doi:10.1155/2019/1651724

25. Dunkelberger JR, Song W-C. Complement and its role in innate and adaptive immune responses. *Cell Res*. 2010;20(1):34-50. doi:10.1038/cr.2009.139

26. Pahwa R, Goyal A, Bansal P, Jialal I. Chronic Inflammation. In: *StatPearls*. StatPearls Publishing; 2020. Accessed July 21, 2020. http://www.ncbi.nlm.nih.gov/books/NBK493173/

27. Ramos PS, Brown EE, Kimberly RP, Langefeld CD. Genetic Factors Predisposing to Systemic Lupus Erythematosus and Lupus Nephritis. *Semin Nephrol*. 2010;30(2):164-176. doi:10.1016/j.semnephrol.2010.01.007

28. Ghodke-Puranik Y, Niewold TB. Immunogenetics of Systemic Lupus Erythematosus: A Comprehensive Review. *J Autoimmun*. 2015;64:125-136. doi:10.1016/j.jaut.2015.08.004

29. Moulton VR, Suarez-Fueyo A, Meidan E, Li H, Mizui M, Tsokos GC. Pathogenesis of Human Systemic Lupus Erythematosus: A Cellular Perspective. *Trends Mol Med*. 2017;23(7):615-635. doi:10.1016/j.molmed.2017.05.006

30. Pan Q, Chen J, Guo L, et al. Mechanistic insights into environmental and genetic risk factors for systemic lupus erythematosus. *Am J Transl Res*. 2019;11(3):1241-1254.

31. Liao JB. Viruses and Human Cancer. *Yale J Biol Med*. 2006;79(3-4):115-122.

32. Smatti MK, Cyprian FS, Nasrallah GK, Al Thani AA, Almishal RO, Yassine HM. Viruses and Autoimmunity: A Review on the Potential Interaction and Molecular Mechanisms. *Viruses*. 2019;11(8). doi:10.3390/v11080762

33. Mueller SN, Rouse BT. Immune responses to viruses. In: Rich RR, Fleisher TA, Shearer WT, Schroeder HW, Frew AJ, Weyand CM, eds. *Clinical Immunology (Third Edition)*. Mosby; 2008:421-431. doi:10.1016/B978-0-323-04404-2.10027-2

34. Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D, Darnell J. Viruses: Structure, Function, and Uses. *Mol Cell Biol 4th Ed*. Published online 2000. Accessed September 15, 2020. https://www.ncbi.nlm.nih.gov/books/NBK21523/

35. Coico R, Sunshine G. *Immunology: A Short Course*. Seventh edition. John Wiley & Sons Inc; 2015.

36. Rosendahl Huber S, van Beek J, de Jonge J, Luytjes W, van Baarle D. T cell responses to viral infections - opportunities for Peptide vaccination. *Front Immunol*. 2014;5:171. doi:10.3389/fimmu.2014.00171

37. Charles A Janeway J, Travers P, Walport M, Shlomchik MJ. B-cell activation by armed helper T cells. *Immunobiol Immune Syst Health Dis 5th Ed*. Published online 2001. Accessed September 15, 2020. https://www.ncbi.nlm.nih.gov/books/NBK27142/

38. Lam JH, Baumgarth N. The Multifaceted B Cell Response to Influenza Virus. *J Immunol Baltim Md 1950*. 2019;202(2):351-359. doi:10.4049/jimmunol.1801208

39. Thaper D, Prabha V. Molecular mimicry: An explanation for autoimmune diseases and infertility. *Scand J Immunol*. 2018;88(2):e12697. doi:10.1111/sji.12697

40. Rojas M, Restrepo-Jiménez P, Monsalve DM, et al. Molecular mimicry and autoimmunity. *J Autoimmun*. 2018;95:100-123. doi:10.1016/j.jaut.2018.10.012

41. Zabriskie JB, Freimer EH. An immunological relationship between the group. A streptococcus and mammalian muscle. *J Exp Med*. 1966;124(4):661-678. doi:10.1084/jem.124.4.661

42. Pacheco Y, Acosta-Ampudia Y, Monsalve DM, Chang C, Gershwin ME, Anaya J-M. Bystander activation and autoimmunity. *J Autoimmun*. 2019;103:102301. doi:10.1016/j.jaut.2019.06.012

43. Kim T-S, Shin E-C. The activation of bystander CD8+ T cells and their roles in viral infection. *Exp Mol Med*. 2019;51(12). doi:10.1038/s12276-019-0316-1

44. Lee H-G, Lee J-U, Kim D-H, Lim S, Kang I, Choi J-M. Pathogenic function of bystander-activated memory-like CD4+ T cells in autoimmune encephalomyelitis. *Nat Commun*. 2019;10. doi:10.1038/s41467-019-08482-w

45. Didona D, Zenzo GD. Humoral Epitope Spreading in Autoimmune Bullous Diseases. *Front Immunol*. 2018;9. doi:10.3389/fimmu.2018.00779

46. Hussein HM, Rahal EA. The role of viral infections in the development of autoimmune diseases. *Crit Rev Microbiol*. 2019;45(4):394-412. doi:10.1080/1040841X.2019.1614904

47. Miller SD, Vanderlugt CL, Begolka WS, et al. Persistent infection with Theiler's virus leads to CNS autoimmunity via epitope spreading. *Nat Med*. 1997;3(10):1133-1136. doi:10.1038/nm1097-1133

48. Boldison J, Khera TK, Copland DA, et al. A novel pathogenic RBP-3 peptide reveals epitope spreading in persistent experimental autoimmune uveoretinitis. *Immunology*. 2015;146(2):301-311. doi:10.1111/imm.12503

49. Steed AL, Stappenbeck TS. Role of viruses and bacteria-virus interactions in autoimmunity. *Curr Opin Immunol*. 2014;31:102-107. doi:10.1016/j.coi.2014.10.006

50. Stölzel U, Schuppan D, Tillmann HL, et al. Autoimmunity and HCV infection in porphyria cutanea tarda: a controlled study. *Cell Mol Biol Noisy--Gd Fr*. 2002;48(1):43-47.

51. Chen J, Zhang H, Chen P, et al. Correlation between systemic lupus erythematosus and cytomegalovirus infection detected by different methods. *Clin Rheumatol*. 2015;34(4):691-698. doi:10.1007/s10067-015-2868-3

52. Woolhouse M, Scott F, Hudson Z, Howey R, Chase-Topping M. Human viruses: discovery and emergence. *Philos Trans R Soc B Biol Sci*. 2012;367(1604):2864-2871. doi:10.1098/rstb.2011.0354

53. Wang L, Wang F-S, Gershwin ME. Human autoimmune diseases: a comprehensive update. *J Intern Med*. 2015;278(4):369-395. doi:10.1111/joim.12395

54. Xu GJ, Kula T, Xu Q, et al. Comprehensive serological profiling of human populations using a synthetic human virome. *Science*. 2015;348(6239):aaa0698. doi:10.1126/science.aaa0698

55. Mohan D, Wansley DL, Sie BM, et al. PhIP-Seq Characterization of Serum Antibodies Using Oligonucleotide Encoded Peptidomes. *Nat Protoc*. 2018;13(9):1958-1978. doi:10.1038/s41596-018-0025-6

56. Sanders MF, Bowman JL. *Genetic Analysis: An Integrated Approach*. Third edition. Pearson Education, Inc; 2019.

57. Churko JM, Mantalas GL, Snyder MP, Wu JC. Overview of High Throughput Sequencing Technologies to Elucidate Molecular Pathways in Cardiovascular Diseases. *Circ Res*. 2013;112(12). doi:10.1161/CIRCRESAHA.113.300939

58. Kosuri S, Church GM. Large-scale de novo DNA synthesis: technologies and applications. *Nat Methods*. 2014;11(5):499-507. doi:10.1038/nmeth.2918

59. Smith GP. Filamentous Fusion Phage: Novel Expression Vectors that Display Cloned Antigens on the Virion Surface. *Science*. 1985;228(4705):1315-1317.

60. Deng X, Wang L, You X, Dai P, Zeng Y. Advances in the T7 phage display system (Review). *Mol Med Rep*. 2018;17(1):714-720. doi:10.3892/mmr.2017.7994

61. Gershoni JM, Roitburd-Berman A, Siman-Tov DD, Freund NT, Weiss Y. Epitope Mapping. *Biodrugs*. 2007;21(3):145-156. doi:10.2165/00063030-200721030-00002

62. Burnham C-AD, McAdam AJ. Your Viral Past: A Comprehensive Method for Serological Profiling to Explore the Human Virome. *Clin Chem*. 2016;62(3):426-427. doi:10.1373/clinchem.2015.245027

63. Kumar VS, Webster M. Exposing the Human Virome. *Clin Chem*. 2015;61(10):1311-1312. doi:10.1373/clinchem.2015.246462

64. Burbelo PD, Iadarola MJ, Chaturvedi A. Emerging technologies for the detection of viral infections. *Future Virol*. 2019;14(1):39-49. doi:10.2217/fvl-2018-0145

65. Mina MJ, Kula T, Leng Y, et al. Measles virus infection diminishes preexisting antibodies that offer protection from other pathogens. *Science*. 2019;366(6465):599-606. doi:10.1126/science.aay6485

66. Bender Ignacio RA, Dasgupta S, Stevens-Ayers T, et al. Comprehensive viromewide antibody responses by systematic epitope scanning after hematopoietic cell transplantation. *Blood*. 2019;134(6):503-514. doi:10.1182/blood.2019897405

67. Hill JA, Krantz EM, Hay KA, et al. Durable preservation of antiviral antibodies after CD19-directed chimeric antigen receptor T-cell immunotherapy. *Blood Adv*. 2019;3(22):3590-3601. doi:10.1182/bloodadvances.2019000717

68. Isnard P, Kula T, Avettand Fenoel V, et al. Temporal virus serological profiling of kidney graft recipients using VirScan. *Proc Natl Acad Sci U S A*. 2019;116(22):10899-10904. doi:10.1073/pnas.1821166116

69. Johnson TP, Larman HB, Lee M-H, et al. Chronic Dengue Virus Panencephalitis in a Patient with Progressive Dementia with Extrapyramidal Features. *Ann Neurol*. 2019;86(5):695-703. doi:10.1002/ana.25588

70. Leon KE, Schubert RD, Casas-Alba D, et al. Genomic and serologic characterization of enterovirus A71 brainstem encephalitis. *Neurol Neuroimmunol Neuroinflammation*. 2020;7(3). doi:10.1212/NXI.0000000000000703

71. Schubert RD, Hawes IA, Ramachandran PS, et al. Pan-viral serology implicates enteroviruses in acute flaccid myelitis. *Nat Med*. 2019;25(11):1748-1752. doi:10.1038/s41591-019-0613-1