Promoter Polymorphisms in Interferon Regulatory Factor 5

Daniel N. Clark
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Microbiology Commons

BYU ScholarsArchive Citation
Clark, Daniel N., "Promoter Polymorphisms in Interferon Regulatory Factor 5" (2013). Theses and Dissertations. 4058.
https://scholarsarchive.byu.edu/etd/4058

This Dissertation is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
Promoter Polymorphisms in Interferon Regulatory Factor 5

Daniel N. Clark

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Brian Douglas Poole, Chair
Greg L. Burton
F. Brent Johnson
Steven M. Johnson
Kim L. O’Neill

Department of Microbiology and Molecular Biology
Brigham Young University
June 2013

Copyright © 2013 Daniel N. Clark
All Rights Reserved
ABSTRACT

Promoter Polymorphisms in Interferon Regulatory Factor 5

Daniel N. Clark
Department of Microbiology and Molecular Biology, BYU
Doctor of Philosophy

The promoter region of interferon regulatory factor 5 (IRF5) contains the rs2004640 T or G single nucleotide polymorphism (SNP) and a CGGGG indel. Both of these polymorphisms have been implicated as genetic risk factors for several autoimmune diseases, including systemic lupus erythematosus, whose pathology involves altered apoptosis and cytokine signaling. The polymorphisms’ overall effect is to increase IRF5 levels. IRF5 is a transcription factor of several cytokines, including interferon, and is pro-apoptotic. Thus an alteration of cytokine levels and apoptosis signaling due to high IRF5 levels is the proposed source of autoimmune risk.

Each of IRF5’s four first exons (1A, 1B, 1C, 1D) has its own promoter and responds to specific stimuli. rs2004640 is a T or G polymorphism; T is the risk allele. The SNP creates a sequence-specific recognition site for the spliceosome, making exon 1B spliceable. Analysis of the 1B promoter showed putative p53 binding site. IRF5 and p53 are pro-apoptotic transcription factors, and the p53 site may provide a positive feedback loop. Apoptosis levels were altered in cells with the rs2004640 risk T/T allele when treated with DNA damaging agents (extrinsic apoptosis), but not when activating death receptors (intrinsic apoptosis). The 1B promoter was the only one to activate expression after inducing DNA damage in a luciferase reporter assay, and this activation was abolished after mutating the p53 site. The exon 1A promoter contains either three or four copies (4X) of CGGGG; the 4X variant is the risk allele. The 1A promoter is constitutively active and is responsive to the Toll-like receptor 7 agonist imiquimod.

RNA folding analysis revealed a hairpin encompassing exon 1B. Mutational analysis showed that the hairpin shape decreased translation five-fold in a luciferase reporter assay. Cells with the CGGGG or rs2004640 risk allele exhibited higher levels of IRF5 mRNA and protein, but demonstrated no change in mRNA stability. Quantitative PCR in cell lines with either risk polymorphism demonstrated decreased usage of exons 1C or 1D, although no other correlated splicing events were observed. Also, several mRNA splice variants of IRF5 were sequenced.

The risk polymorphisms altered cytokine signaling as well. Expression of interferon, Toll-like receptor, and B cell receptor pathways were affected by a risk haplotype which includes the rs2004640 SNP. The CGGGG polymorphism decreased the levels of CC-chemokine receptor 7.

Specific transcription factor binding sites define promoter activity and thus first exon usage and transcription levels. Translation levels are affected by mRNA folding. Overall, the rs2004640 SNP and the CGGGG indel cause high levels of IRF5. High IRF5 expression causes altered cytokine and apoptosis signaling, and may bias the immune system toward autoimmunity.

Keywords: autoimmunity, cytokines, apoptosis
ACKNOWLEDGEMENTS

I wish to acknowledge and thank my wife. Her support was essential to beginning and finishing this large project, as well as the day-to-day work. Without Rachel Clark, this could never have happened.

Thanks go to Dr. Brian Poole, my adviser. His advice, support, and guidance were vital to the entire process. The feeling that he had my back was a true comfort.

Drs. Gregory Burton, Brent Johnson, Steve Johnson, and Kim O’Neill were amazing committee members who helped with every need. Thanks also go to Drs. Laura Bridgewater and Chin-Yo Lin, who served as members during several meetings.

Many undergraduate researchers became my helpers, colleagues, and friends. They will be remembered.

Many professors, many graduate students, and many staff members in Brigham Young University’s Microbiology and Molecular Biology department and all across campus were instrumental in their help.
**Table of contents**

Title page ......................................................................................................................................... i

ABSTRACT .................................................................................................................................... ii

ACKNOWLEDGEMENTS ........................................................................................................... iii

Table of contents ............................................................................................................................ iv

List of tables ................................................................................................................................... viii

List of figures ................................................................................................................................... ix

1. Introduction and literature review ............................................................................................... 1

1.1. Systemic lupus erythematosus ............................................................................................. 1

1.2. Genetic risk factors for SLE ................................................................................................ 1

1.3. Environmental risk factors for SLE ..................................................................................... 3

1.4. Function of interferon in the context of SLE ....................................................................... 4

1.5. Apoptosis due to interferon; the connection between SLE and apoptosis ......................... 7

1.6. Mouse models for the study of IFN and apoptosis pathways .............................................. 8

1.7. A cycle of autoantibody production ................................................................................... 10

1.8. Clinical component interferon and SLE ............................................................................ 12

1.8.1. Clinical trials of anti-cytokine inhibitors .................................................................... 14

1.9. Interferon signaling pathway genes identified in SLE genetic risk screens ................. 15

1.9.1. Interferon production pathways .................................................................................. 18
2.6. Altered expression levels of associated genes due to the rs2004640 and CGGGG risk factors........................................................................................................................................ 60

3. Summary and discussion........................................................................................................... 66

3.1. Promoter differences .......................................................................................................... 66

3.2. Apoptosis differences ......................................................................................................... 70

3.3. mRNA differences ............................................................................................................. 71

3.4. Downstream expression differences .................................................................................. 74

3.5. Conclusion ......................................................................................................................... 76

4. Material & Methods .................................................................................................................. 78

4.1. Plasmid construction and luciferase assay ......................................................................... 78

4.2. Cell lines ............................................................................................................................ 79

4.3. Genotyping of volunteers and formation of paired samples .............................................. 79

4.4. Cell treatments ................................................................................................................... 79

4.5. cDNA libraries and PCR .................................................................................................... 80

4.6. Sequencing ......................................................................................................................... 83

4.7. Statistical analysis ............................................................................................................. 83

4.8. Computer Programs ........................................................................................................... 84

4.9. Western blotting ................................................................................................................. 84

4.10. Enzyme-linked immunosorbent assay (ELISA) .............................................................. 85

4.11. Information about microarray data .................................................................................. 85
Appendix A. Effect of imiquimod ........................................................................................................... 86

Appendix B. Herpes simplex virus typing assay affected by region-specific phylogenetic changes
............................................................................................................................................................... 88

Appendix C. Review of interferon alpha in systemic lupus erythematosus ................................. 93

Appendix D. Review of cytokine inhibitors for the treatment of systemic lupus erythematosus
.............................................................................................................................................................................. 102

Appendix E. Review of interferon and apoptosis in systemic lupus erythematosus .............. 112

Appendix F. Microarray analysis of B cell expression differences due to an rs2004640 risk haplotype................................................................................................................................. 133

5. References.................................................................................................................................................. 146
List of tables

Table 1. Description of the interferon response signature .............................................................. 7
Table 2. Description of genome wide association studies .................................................................. 17
Table 3. Genes exhibiting differential expression between risk and protective cells in the canonical pathways identified through Ingenuity Pathway Analysis ............................................ 60
Table 4. List of primers and oligonucleotide sequences, and PCR and annealing conditions ..... 81
List of figures

Figure 1. Interferon protein structures ............................................................................................ 4
Figure 2. Cell-to-cell interferon signaling and its effects ............................................................... 6
Figure 3. Production of interferon due to defects in apoptosis signaling ....................................... 8
Figure 4. The cycle of altered immune response in SLE .............................................................. 11
Figure 5. Potential lupus drugs and the pathways they affect ........................................................ 15
Figure 6. Lupus risk genes that affect interferon production pathways ....................................... 19
Figure 7. Genes associated with risk for lupus in the apoptosis pathway .................................. 22
Figure 8. IRF5 mRNA and protein structures ............................................................................. 26
Figure 9. IRF5 mRNA and the position of the rs2004640 SNP and CGGGG indel .................... 28
Figure 10. Functional changes due to IRF5 promoter polymorphisms ......................................... 30
Figure 11. Demographics of samples collected for testing .......................................................... 34
Figure 12. Promoter analysis results for each first exon of IRF5 ................................................. 36
Figure 13. Transcriptional activity of each promoter of IRF5 ..................................................... 38
Figure 14. Activation of IRF5 transcription through exon 1A and 1D promoters upon imiquimod treatment .............................................................................................................. 40
Figure 15. Inverted repeat in the promoter region of IRF5 ......................................................... 41
Figure 16. p53-dependent activity of IRF5’s 1B promoter upon DNA damage ......................... 43
Figure 17. Altered apoptosis levels due to the rs2004640 polymorphism in IRF5 ...................... 46
Figure 18. Altered intrinsic but not extrinsic apoptosis levels in rs2004640 risk T/T cells ............ 47
Figure 19. Doubled IRF5 mRNA and protein levels in cells with the rs2004640 or CGGGG risk polymorphisms .............................................................................................................. 49
Figure 20. IRF5 mRNA stability not affected by the rs2004640 or CGGGG risk polymorphisms
.................................................................................................................................................. 50

Figure 21. Decreased usage of exon 1C and 1D and increased 1B in cells with the rs2004640 risk
 allele; decreased exon 1D usage in cells with the CGGGG risk allele................................. 52

Figure 22. Different translational activity levels among the four 5’UTRs of IRF5 first exons.... 55

Figure 23. 5’ hairpin in exon 1B-containing IRF5 mRNA............................................................. 56

Figure 24. Inhibition of exon 1B translation efficiency by a hairpin structure......................... 57

Figure 25. IRF5 mRNA variants organized by splicing events, including five novel variants .... 59

Figure 26. Altered expression of genes in the interferon, Toll-like receptor and B cell receptor
 pathways due to an IRF5 risk haplotype.................................................................................63

Figure 26 caption.......................................................................................................................64

Figure 27. Decreased CCR7 expression in LCLs with the IRF5 CGGGG indel 4X/4X genotype
 by quantitative PCR ..................................................................................................................65

Figure 28. Altered gene expression due to imiquimod treatment by PCR or ELISA............ 87
1. Introduction and literature review

1.1. Systemic lupus erythematosus

Systemic lupus erythematosus (SLE) is the prototypic disease for the study of systemic autoimmunity. It involves the production of anti-nuclear antibodies (ANA). Since the immune system recognizes components of the nucleus as foreign, the symptoms can be extremely varied. Some of these include skin rash and sun sensitivity, anemia, fatigue, joint pain, kidney and liver damage, and neurological symptoms such as headache and depression. Lupus is about ten times more common among women. The treatment is directed against the symptoms—as opposed to the causes—of the disease. These treatments are usually immunosuppressants. With current medical interventions, the disease is rarely fatal in developed countries.

SLE is generally diagnosed long after the disease begins. This means that the cause of the disease is hard to find, buried in the past. In the search for the elusive causal agents for SLE, one candidate is the immune signaling molecule, interferon (IFN). Interferon is a secreted signaling protein, or cytokine, which is expressed at higher levels in SLE patients and has been associated with incidence and severity of the disease.

1.2. Genetic risk factors for SLE

A combination of environmental triggers and genetic susceptibility combine to initiate SLE. Support for a genetic component to the disease includes a high sibling risk ratio (λS between 5.8 and 29), high heritability (greater than 66%), and higher concordance rates between monozygotic twins (20 to 40%) compared to other full siblings and dizygotic twins (2 to 5%) [1-3]. A large number of genetic risk factors are associated with increased susceptibility to the SLE. This
genetically determined increased risk status has been referred to as a threshold liability [4], which is a polygenic state that varies between individuals. Environmental factors also affect lupus susceptibility and interact with this threshold liability, but just as is the case for genetic factors, there is no single environmental cause. An individual may have only a few of the genetic risk variations (low threshold liability) and never get SLE despite exposure to environmental triggers. In contrast, another person may have many of these variations (high threshold liability) and then develop SLE on first exposure to an environmental trigger.

Although there are many etiological components, they usually converge on a heightened state of activation for the immune system, with resultant increases in interferon and other cytokine production and signaling. That is to say that immune dysregulation could be thought of as either a causative agent, a result of the disease, or both.

This dissertation will discuss a transcription factor called interferon regulatory factor 5 (IRF5). IRF5 is a direct target for interferon and other cytokine signaling and is proapoptotic. We will begin with an introduction to the basics of interferon function, and how dysregulation of apoptosis can lead to interferon production due to immune complexes. We will then discuss how the functioning of the immune system changes in someone with SLE, other genes which are associated with risk for SLE, and clinical aspects of interferon in SLE, including anticytokine therapies. Finally, we will discuss two genetic risk polymorphisms for lupus: the T allele of the rs2004640 single nucleotide polymorphism (SNP) and the four-copy variant of the CGGGGG indel. These polymorphisms are within the IRF5 gene, and their effect is the focus of this study.
1.3. Environmental risk factors for SLE

Several environmental triggers have been noted for lupus. They include drug treatments such as procainamide, isoniazid, and interferon treatments. Although upon discontinuation of treatment, the symptoms are generally reversed. Others include UV light exposure and infectious agents such as viruses.

Epstein-Barr virus (EBV) infection is associated with lupus, likely a causal relationship. Lupus patients have increased EBV seroprevalence and elevated serum titers of anti-EBV antibodies [5]. Among those with SLE, there is also an elevation of anti-EBV antibodies which precedes lupus flares [5]. EBV infection is also not controlled in lupus patients as well as it is in controls, including up to forty-fold increased viral load and an altered T-cell response [6-7].

EBV antigens such as EBV nuclear antigen 1 (EBNA1) are targets of antibodies for the immune system. Once the body makes these antibodies, a similarity in protein structure can create a cross reactivity with self proteins, such as Sm and RNP, which are important lupus autoantigens [8-13]. Since lupus is diagnosed long after the causal events such as antibody cross-reactivity have occurred, and since EBV infection is nearly ubiquitous in adults, an association of lupus to EBV positivity is not necessarily a strong case for association. However, lupus has been associated with prior EBV infection in not just adult lupus patients, but also in pediatric cases [14-20]. Since EBV positivity is much lower in pediatric cases, the higher rates of lupus among EBV-positive pediatric cases is more striking. These changes in viral infection or the response to viral infection such as interferon or other cytokine levels may be due to the genetic risk factors mentioned above (Section 1.2), or simply a symptom of the disease.
1.4. Function of interferon in the context of SLE

Interferon is a signaling protein which is secreted to activate neighboring cells in response to viruses, other infections, or as part of various immune signaling pathways. It is a cytokine, or secreted immune signaling protein which allows communication between cells. When a cell is infected with a virus, interferon is produced and secreted as a warning to other cells to prepare for an infection. Interferons alpha (IFNα) and beta (IFNβ) are the type I interferons, and interferon gamma (IFNγ) is the type II interferon (Figure 1). Most of the cells in the human body have receptors for type I IFN, whereas only certain immune cells express the receptor for type II IFN [21]. The proteins are made by many different cells, but generally speaking, IFNα is of leukocyte origin, IFNβ is of fibroblast origin, and IFNγ is made by lymphocytes [22]. Other less studied interferons also exist, and interferons are conserved among many species. Both alpha and beta types signal through the IFNα receptor (IFNAR), and then through Jak/STAT signaling.

![Interferon protein structures](image)

**Figure 1. Interferon protein structures.** Interferons alpha and beta, the type I interferons, have a common structure composed mainly of five alpha helices. Shown are IFNα2a and IFNβ1 monomers based on protein data bank (PDB) files 1itf and 1au1, respectively. Although the monomers of each are very similar in structure, the functional form of both is a dimer, and the two dimerize differently; IFNα2a along homologous surfaces and IFNβ1 on opposing sides of the protein [23]. IFNγ is show in its dimerized form, with the two colors representing two intertwined monomers, based on PDB file 1hig. Not shown to scale; figures drawn with Jmol [24].
pathways to alter gene expression. IFNα requires the Jak family member tyrosine kinase 2 (Tyk2), and it thus sends a similar though different signal than IFNβ [25]. This review portion (Section 1) will discuss type I interferons, IFNα and IFNβ, which herein may be referred to simply as interferon. When interferon gamma (IFNγ) is discussed, it will be noted explicitly.

The main purpose of interferon is to shut down a cell before a virus can take it over, although it has many other jobs [26]. Interferon signaling leads to increased apoptosis, which is a normal response to control viral spread or to decrease the size of a tumor [27]. If one cell can undergo apoptosis before a virus can replicate and infect other cells, the infection is halted [28].

Interferon can be produced in response to infection, other cytokines, mitogens, and several signaling pathways. Once produced it is secreted where it can be recognized by other cells, which is called paracrine signaling, or by the cell which produced it, called autocrine signaling. One type of cell, the plasmacytoid dendritic cell (pDC), is a natural interferon producer, and its ability to make very large amounts of IFNα is incriminated in the pathogenesis of lupus [29].

Several effects occur when interferon ligates an interferon receptor (Figure 2). Jak/STAT signaling pathways are activated which alters gene expression [30]. Interferon causes an increase in the expression of both of the major histocompatibility complexes (MHCI and MHCII) for presentation of viral peptides to T cells, which can then lead to activation of other cells in order to kill infected cells and remove them [31]. Interferon also increases intracellular levels of protein kinase R (PKR) which recognizes viral nucleic acids and activates RNase L to degrade viral RNAs. PKR also slows protein synthesis by inactivating translational initiation factors, so
that viral protein synthesis is slowed [32]. p53 is also activated, which is pro-apoptotic [27]. Interferons activate immune cells, especially natural killer cells and macrophages [33].

Figure 2. Cell-to-cell interferon signaling and its effects. One cell produces interferon and either another cell (paracrine signaling) or the same cell (autocrine signaling) receives the signal. Cellular processes that are activated or otherwise altered are indicated in the target cell. ↑: an increase, ↓: a decrease, MHC: major histocompatibility complex, PKR: protein kinase R, NK: natural killer cell, MΦ: macrophage

This activation cascade is normally turned off after an infection is cleared to prevent damage to uninfected cells. However this activation state is not reduced to the normal levels in individuals with SLE, where a higher level of interferon is present [34-35]. This higher amount of interferon is also measurable by an increase in the expression of interferon-stimulated genes seen in lupus patients, called the interferon response signature [36-38] (See Table 1). When IFN is turned on it actively affects how other cells are functioning.

As a general feature of autoimmune diseases such as SLE, the immune system is in an “always on” state, which can lead to a breach in the body’s natural tolerance to self. Once this self tolerance is lost, autoimmune disease can result. In addressing why the immune system generates
an attack against one’s own body, the over-activation of the immune system—including the overproduction of interferon in SLE patients—is a part of this picture.

Table 1. Description of the interferon response signature. The effect of interferon can be measured by the changes it produces in downstream gene expression. Lupus patients have an altered interferon response signature.

<table>
<thead>
<tr>
<th>Interferon Response Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Defined as the list of genes that are upregulated in peripheral blood of many SLE patients</td>
</tr>
<tr>
<td>• Identified using microarray analysis [37]</td>
</tr>
<tr>
<td>• Genes whose expression levels change in response to interferon [39]</td>
</tr>
<tr>
<td>• Confirmed using multiple array systems and RT-PCR [37-40]</td>
</tr>
<tr>
<td>• SLE patient clusters can be generated according to IFN response signature [38, 41]</td>
</tr>
</tbody>
</table>

1.5. Apoptosis due to interferon; the connection between SLE and apoptosis

One effect of interferon production is the release of autoantigens due to increased cell death. This release is normally controlled by a process called efferocytosis [42], or apoptotic cell removal, where cell debris are processed by immune cells or neighboring cells which remove them by phagocytosis (Figure 3). Defects in apoptotic pathways have been noted in individuals with SLE [43]. Instances of why this occurs have been studied. For example, in SLE patients there is an overexpression of both soluble and membrane-bound Fas. Fas is a receptor which signals a cell to undergo apoptosis when complexed with its receptor, Fas ligand (Fas L). The levels of Fas also correlate with the amount of apoptotic lymphocytes and disease activity of SLE [44-45]. Mouse models of lupus commonly have genetic variations in apoptotic pathways such the Fas/Fas L pathway and interferon pathways.
Mouse lupus models as well as human SLE patients make antibodies to self antigens. This is likely because of over-exposure of potential autoantigens to the immune system. This could be due to an increased amount of apoptosis or a decrease in the rate of clearance of apoptotic debris. Apoptosis can be induced by interferon, Fas, and other signaling events. However, apoptosis is also part of the natural cycle of cellular growth and death for every cell type. Cells undergoing apoptosis are recognized as dead by other cells, so that they are cleared [46].

**Figure 3. Production of interferon due to defects in apoptosis signaling.** Apoptotic cells form naturally or due to damage. If the apoptotic debris is cleared normally, the debris is efferocytosed by neighboring or immune cells such as macrophages, and there is no immune response. In SLE, apoptotic debris remains present for the immune system to recognize. This can be due to either an increase in apoptosis or a decrease in clearance of apoptotic debris. If contents are released, they can form immune complexes with autoantibodies. These immune complexes can cause cells to produce interferon.

1.6. Mouse models for the study of IFN and apoptosis pathways

Mouse models have been very useful in understanding the etiology and pathogenesis of lupus. Two approaches to experimental mice have been used to generate information about the role of interferon in lupus. In the first approach, autoimmune-related genes are knocked out and the resulting effects on lupus are studied. For the second, established lupus mouse models are
studied on a molecular level for differences in autoimmune pathways such as interferon-related effects. These two approaches often overlap, as in cases where interferon-related genes are knocked out in lupus-prone mice. Several established lupus mouse models include the Murphy Roths Large/lymphoproliferative (MRL/lpr) mice, New Zealand black/New Zealand white (NZB/NZW), and others. These are mice that spontaneously develop lupus, and several of them have been investigated to understand the role of interferon in their pathogenesis. The following are a few illustrative examples which represent the power of these model systems.

One mouse model that is especially relevant for the study of interferon in lupus is the Black 6 cross to satin beige/Murphy group of the Jackson laboratory (BXSB/MpJ) or Y-linked autoimmune accelerator (Yaa) mouse. These mice spontaneously develop a lupus-like disease in a sex-linked fashion because of a duplication of the Toll-like receptor 7 (TLR7) gene on the Y chromosome [47]. TLR7 is responsible for inducing interferon in response to viral infection or autoantibody production.

Another interesting mouse for the study of interferon is the NZB/NZW mouse. These mice spontaneously develop a lupus-like autoimmune disease. They have been used to investigate the role of several interferon-related molecules and cells. For example, treating these mice with interferon accelerates disease in a T-cell dependent manner [48-49], while knocking out or inhibiting interferon-related genes slows or eliminates the development of lupus-like symptoms [50-51]. These mice have been used to clarify the interactions between sex hormones and interferon in lupus etiology [52-54], and they serve as an excellent all-around model for spontaneous development of lupus.
The role of several interferon-related molecules has been examined using a combination of mouse models. As an example, consider the gene interferon regulatory factor five (IRF5). This gene is an interferon-regulating gene which will be described in Section 1.10 below. It was discovered that knocking out *Irf5* prevents or inhibits the development of lupus in MRL/lpr mice, Fcγ−/− Yaa mice, and pristane-injected mice [55-57].

Mouse models for lupus represent a powerful and flexible mechanism for investigating the role of multiple aspects of lupus. However, it must be remembered that the mutations or disease manifestations in these mice are not necessarily related to those seen in human lupus, and therefore the results observed must be interpreted with caution.

### 1.7. A cycle of autoantibody production

When it comes to SLE we may think of interferon production as a cycle which begins when an environmental trigger such as a viral infection, UV light damage, or medical treatment activates the immune system to produce interferon.

Normally B cells, which produce antibodies to self-antigens, undergo negative selection where they receive signals to die off or become inactivated if they make antibody against a self antigen. This self tolerance is breached in SLE [58], and the self antigens released from damaged or apoptotic cells during or after initial triggering events become the targets of autoantibodies. When autoantibodies are produced, they are either made by B cells or plasma cells. Plasma cells are a mature differentiated form of B cells, which secrete antibodies instead of maintaining them bound to the cell surface.
Autoantibodies stimulate the production of interferon when they form immune complexes, which are immunostimulatory [59]. Immune complexes are composed of aggregates of antibody and antigen molecules. Since immune complexes must processed by the body, they are a main source of SLE pathology, as they obstruct small passages in areas of the body such as kidney glomeruli and joints [60]. Immune complexes are a main cause of end organ damage and therefore mortality in lupus. However, lupus is less fatal than it once was due to effective treatment of symptoms.

![Figure 4. The cycle of altered immune response in SLE.](image)

As part of the normal immune response, the presence of an antigen results immune activation; this may include interferon production. In blue is a cycle which exists in SLE, amplifying the amount of IFN produced. This can happen due to alterations in the immune response. If autoantibodies form due to lowered self tolerance or increased apoptosis, immune complexes may form which are themselves immunostimulatory. Once this cycle becomes sustained, it can leave the immune system in an “always on” state.

Immune complexes may include the common SLE autoantigens such as RNA-containing protein complexes like Sm, RNPs, Ro, and La. These are of nuclear origin and have a combination of both nucleic acids and protein [61]. Once complexed with antibody, this combination of molecule types means many pathways can be turned on. For example, antibody can stimulate an
immune cell through an Fc receptor, nucleic acids can stimulate cells through Toll-like receptors, and proteins can be recognized by other antibodies.

Immune cells are activated by immune complexes and the cycle continues (Figure 4). Interferon production is instigated by stimulated immune cells which recognize part of the complex, be it the antibody, the antigen, or other associated molecules.

1.8. Clinical component interferon and SLE

The body of evidence in this section will describe the clinical data which associate interferon to systemic lupus erythematosus. Many researchers have sought to determine if higher levels of IFN, which is common in lupus patients, is a cause of lupus or an effect of lupus.

Intensity of lupus is scored by several different methods: the SLE disease activity index (SLEDAI), a common modification of SLEDAI called Safety of Estrogens in Lupus Erythematosus National Assessment (SELENA), and the British Isles Lupus Assessment Group (BILAG) index. These measure many clinical indicators on a weighted scale. Clinical signs and symptoms evaluated include seizure, psychosis, organic brain syndrome, visual disturbance, cranial nerve disorder, lupus headache, cerebrovascular accidents, vasculitis, arthritis, myositis, urinary casts, hematuria, proteinuria, pyuria, new rashes, alopecia, mucosal ulcers, pleurisy, pericarditis, low complement levels, increased DNA binding, fever, thrombocytopenia, and leucopenia [62].

12
An interesting occurrence can happen when someone undergoes treatment with IFNα. The presence of increased levels of IFN leads to lupus or a lupus-like syndrome [63-65]. Because the lupus symptoms usually disappear after IFN treatment ends, this connection suggests that IFN may be more of a cause than an effect. In a small number of cases, some patients also develop SLE as a result of these IFN treatments, and in these cases the IFN is also the disease trigger. Furthermore, within a family, the levels of interferon among all members correlate, suggesting that this is a heritable trait [66]. That is, even the siblings of a lupus patient with high IFN levels are more likely to have higher IFN levels. This also supports a causal role for IFN. The causal role does not preclude the role of interferon as an effect of the disease as well.

Clinically, disease activity can be measured and correlated to other observations to determine the cause of the different levels of activity. One item linked to SLE activity is IFN, where higher levels of IFN in the serum correlated with more severe disease in most cases [38, 67-71].

Common lupus autoantibodies also correlate with IFN levels. A very strong correlation is consistently observed between IFNα levels and the presence of antibodies directed against common SLE autoantigens like Ro, La, Sm, RNP, and dsDNA [72].

Another set of findings has to do with properties of the main producer of IFN, plasmacytoid dendritic cells (pDCs). High numbers of IFN-producing pDCs have been observed in lupus skin lesions [73-74]. Since the cells are present at the scene of the crime, the increased interferon could have to do with the pathology in these cases.
1.8.1. Clinical trials of anti-cytokine inhibitors

At the time of writing, two anti-interferon clinical drug trials for SLE are being conducted. They are for Sifalimumab and Rontalizumab, which are both anti-interferon monoclonal antibodies in phase II trials [75-76]. Another phase II trial is underway, whose drug MEDI-546 targets the interferon-α receptor 1 (IFNAR1) [77]. All of these monoclonal antibodies are designed to block interferon alpha signaling by preventing its recognition by neighboring cells. If these drugs are found to be effective by reducing SLEDAI or BILAG scores, it will further show that IFN plays a critical role in the pathogenesis of lupus.

Of note, the United States Food and Drug Administration recently approved an antibody to B lymphocyte stimulator (BLyS, also known as B cell activating factor (BAFF) ) to treat SLE [78]. This drug, called Belimumab, is the first new lupus drug in around fifty years and should help control B cell activation, selective apoptosis, and autoantibody production to some degree.

Other anti-cytokine therapies have been or likely will be tried. Several of these target the interleukins, a group of important immune cytokines. An example is interleukin-6 (IL-6). IL-6 levels are higher among lupus patients [79]. Furthermore, the IL-6 which B cells make contributes to the production of autoantibodies [80]. Treatment with an antibody directed against the IL-6 receptor, called Tocilizumab, led to decreased SELENA-SLEDAI scores in one study [81]. This and other cytokine inhibitor drugs have undergone testing in mouse models of lupus or human lupus patients (Figure 5). The group of anti-cytokine—including anti-interferon—antibodies currently being studied will likely become the next generation of lupus drugs. Unlike most current treatments, they target the cause of the disease instead of just symptoms.
Figure 5. Potential lupus drugs and the pathways they affect. Several pathways involved in lupus pathogenesis (gray boxes) are mediated by cytokines (white boxes, each represented by a symbol) which are produced by many of the body’s cells as part of normal signaling or during the course of an infection. These pathways are blocked by the clinical administration of monoclonal antibodies, or other antibody fusion proteins, whose names are listed according to their cytokine target. BAFF: B cell activating factor, BP: binding protein, IFN: interferon, IL: interleukin, RA: receptor antagonist, T: Blocks or inhibits.

1.9. Interferon signaling pathway genes identified in SLE genetic risk screens

We have looked at the disease state of SLE and how the immune system functions improperly to instigate disease, and we will now look at genetic risk factors. Pathogenesis begins when an environmental trigger works on the genetic background of varying degrees of susceptibility or threshold liability. Genetic susceptibility is thought to account for at least 20% of the risk for SLE [2]. To find the actual genes involved, studies are performed to determine the linkage or association of a variation in the genome to a particular disease.
One important method is called a genome wide association study (GWAS). These GWA studies genotype thousands of individuals, grouped into SLE patients and healthy controls, comparing them at thousands of single nucleotide polymorphisms (SNPs). Other variations, such as insertion-deletion sequences (indels) and repeat sequences, can also be screened for association. These studies reveal the genomic regions which contain disease-associated genes, because the variations are more common in people with the disease. Individual genes or gene pathways are pinpointed and can ultimately lead to treatment strategies. Many genes have been identified that contain SNPs and other small variations which confer risk to SLE; they will be referred to as risk genes in the following sections.

GWA studies are especially useful for diseases with unknown or complex genetic components. The genome is examined for sets of single nucleotide polymorphisms (SNPs). When sets of SNPs are usually inherited together in a group—and thus statistically associated—it is called a haplotype. When a polymorphism or haplotype is more common in the disease group than in the unaffected group, it can be assumed that it is associated with the disease.

Although specific genes are sometimes found which may predict a disease, it is more likely that the information will reveal molecular pathways associated with the disease. Association of genes or pathways to diseases such as heart disease, asthma, diabetes, and others have been found using this method [82]. The amount of effect is measured as an odds ratio (OR), which is a measure of the strength of association of the disease with a haplotype. A median OR value is around 1.3, with more causal genes having much higher association ORs. For example, one of the lupus-
associated haplotypes TREX1, has a published OR of 25 [83]. In such cases, the genetic risk is almost certainly associated with the disease.

Table 2. Description of genome wide association studies. In order to determine the genetic source of a polygenic trait or disease, GWAS may be performed. Variations in the genome can be associated to the disease or trait being studied.

<table>
<thead>
<tr>
<th>Genome-wide association studies (GWAS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>After completing the Human Genome Project, a logical next step is to find what parts of the genome are associated with human diseases. Genome-wide association studies aim to discover the genetic risk component of a disease by finding differences in a disease group compared to an unaffected control group. This is especially useful for diseases with unknown and complex genetic components.</td>
</tr>
<tr>
<td>The genome is examined for sets of single nucleotide polymorphisms (SNPs) or other polymorphisms. When a polymorphisms or haplotype is more common in the disease group than in the unaffected group, then it is associated with the disease.</td>
</tr>
<tr>
<td>SNP: single nucleotide polymorphism. These single nucleotide changes are what make one person different from another person and may be clues to finding the genetic risk components of disease.</td>
</tr>
<tr>
<td>Odds ratio: A measure of effect size which describes the strength of association between a disease and a genetic variation. Mean OR values are 1.3. Two polymorphisms described in Section 1.10.1 are the rs2004640 SNP (OR=1.4-1.9, p=5.7<em>10^-7 [84]) and the CGGGG indel (OR=1.4-2.0, p=4.6</em>10^-9 [85]) in the gene IRF5.</td>
</tr>
<tr>
<td>Haplotype: A group of co-inherited genetic variations, such as SNPs, which are grouped together for the purpose of studying larger areas of the genome at one time. Statistical association is calculated for verification. Identifying co-inherited SNPs allows for testing of a single marker SNP instead of each individual polymorphism.</td>
</tr>
</tbody>
</table>

An important caveat to these tests is that they answer the question, “What?” but not the question, “How?” That is, they identify genetic loci which confer risk to SLE, but then further studies are needed to show what functional changes affect people with a risk polymorphism. For most
genes, we do not know what functional role they play. However it is promising to note that the most genes thus far identified in lupus GWAS are within pathways which have been previously associated with lupus.

Several review articles have reviewed the findings of many lupus GWAS with varying degrees of certainty [3, 86-89]. In some cases the indicated susceptibility genes are common in many ethnicities and populations, while others are specific to certain groups. The statistical significance of many of these genes is well established, while others are novel and need to be replicated. For example, an IRF5 SNP rs2070197 has been associated with lupus in Caucasian [90], Hispanic [91], and African American populations [92], but no association was found in a Korean population [93]. Another study used five different ethnic populations to have a more robust result of significance of association to lupus of a CGGGG indel in the IRF5 gene [94]. An important finding is that most of the genes that have been identified in GWA studies can be grouped into functional pathways. The current focus will be on lupus-risk associated genes in two pathways: genes in IFN pathways and genes in apoptosis pathways—those involving clearing of apoptotic cells and clearing of immune complexes.

1.9.1. Interferon production pathways

Intracellular signaling pathways which control interferon production include the production of type I interferons by interferon regulatory factors (IRFs) and the production of type II interferon by STAT4. IRFs are activated by TLRs, which are extracellular or endosomal pattern recognition molecules. TLRs 7, 8, and 9 recognize foreign nucleic acids and are endosomal. Maintaining these TLRs in the endosome instead of the cell surface is an important barrier to too frequent
TLR activation, since self nucleic acid is difficult to distinguish from foreign nucleic acid. Once the nucleic acids are brought into the cells through endocytosis, the TLRs become activated to turn on IRFs (Figure 6). TLR 8 and TLR 9 have both been identified as lupus risk genes [95-96].

Figure 6. Lupus risk genes that affect interferon production pathways. The * represents genes which have been identified as having risk variants associated with lupus. The endosomal TLRs (7, 8, and 9) can bind to autoantigenic nucleic acids (squiggly black lines) and signal through a MyD88 complex which can be affected by association with osteopontin (OPN). If it is not impeded by TNFAIP3, this activates an IRAK signaling complex to phosphorylate IRF5 and IRF7 transcription factors to produce type I IFN. Another signaling pathway which activates interferon production is when IL-12 or IL-23 signals through the IL-12 receptor, and then Tyk2/Jak2 activate the STAT4 transcription factor to produce type II IFN. This signaling pathway is common in T helper cells [97].

TLRs begin a signaling cascade through a MyD88 signaling complex. MyD88 activates another confirmed locus of SLE risk, the gene which encodes IL-1 receptor-associated kinase 1 (IRAK1). In Sle1 and Sle3 mouse models of lupus, IRAK deficiency eliminated most lupus symptoms [98], which highlights the importance of IRAK1. Since this gene is on the X chromosome, it could help explain why lupus is more common among women—who have two X chromosomes. The MyD88 complex can be affected by osteopontin (OPN). It regulates IFNα production in plasmacytoid dendritic cells, which are the body’s main IFNα producer cell [99]. The lupus-risk variant of OPN was tied to high IFNα levels in certain lupus patients [100].
Two interacting proteins are involved in tumor necrosis factor alpha (TNFα) signaling and inflammation: TNFα-induced protein 3 (TNFAIP3) and TNFAIP3-interacting protein 1 (TNIP1). TNFAIP3 and TNIP1 are also lupus risk loci [101-102]. TNFAIP3 encodes the protein A20, which abrogates NF B after an inflammatory response, and lupus-risk variants of this gene are associated with blood and kidney manifestations among lupus patients [103]. TNIP1 interacts with TNFAIP3 as well as affecting several other signal transduction pathways.

Interferon regulatory factors are activated next, further downstream of TLRs; they are transcription factors which travel to the nucleus to bind DNA to initiate transcription. IRF5 binds to a sequence specific region of DNA to induce IFN production. IRF5 polymorphisms have been confirmed as risk factors for SLE among several ethnicities [91-92, 104-106]. Among the main genetic variants within IRF5, there are two copy number indels and several SNPs [107]. The first copy number indel is either three or four copies of a CGGGG repeat [108]; the other is either two or four copies of a 30 bp sequence [107]. Some salient SNPs are rs2004640, rs10954213, and rs10488631. The rs2004640 SNP allows use of an alternate first exon, although this does not change the encoded protein [84]. The SNP rs10954213 creates an early polyadenylation sequence, which yields shorter more stable mRNA [109]. The rs10488631 is downstream of IRF5, 3’ of the poly-A site, and has no known functional change, although its association to lupus is well established [108]. Work has shown that these variants increase the amount of IFN in the presence of SLE autoantibodies [110-111]. The IRF5 risk factors generally act to increase the levels of IRF5; however the increase in IRF5 expression is not entirely due to IRF5’s own polymorphisms [112] which means some trans-acting factors also increase IRF5
levels. Study of the rs2004640 SNP and the CGGGG indel are the main focus of this dissertation; they are introduced in greater detail in Sections 1.10.1 and 1.10.2.

IRF7 is associated with SLE risk by its proximity to SNPs in the IRF7/KIAA1542 locus [113-114]. IRF7 SNPs lead to increased IFNα levels, and they change which autoantigens are targeted in the autoimmune response [110].

Signal transducer and activator of transcription 4 (STAT4) is also associated with risk for SLE. It is a transcription factor which activates genes in proliferation, differentiation, and apoptosis pathways. Two STAT4 SNPs have been examined, rs7574865 increases sensitivity to IFNα [115], and rs3821236 causes STAT4 to be transcribed at higher levels and is additive with IRF5 risk loci so that when both are present, the risk to SLE is multiplied [116-117].

1.9.2. Genes associated with apoptosis and immune complexes

Another set of lupus-risk genes can be placed into a functional group of apoptosis-associated genes. As described earlier, defects in apoptosis can lead to the presence of potential autoantigens. For example, after a cell undergoes apoptosis—if it is not cleared by other cells—its contents may be released. The cellular contents can contain molecules such as nucleic acids, RNA binding proteins, and others which are common lupus autoantigens. If antibodies bind to these antigens, a complex of multiple antibodies and multiple antigens can aggregate. The resultant immune complexes can be broken down through reactions with complement components (Figure 7). Contributing to a lack of immune complex clearance, some complement components are found at low levels in SLE patients [118]. If immune complexes are not broken
down by complement, they reach areas such as the kidneys or joints, which can be damaged by these immune complexes. This is how organ damage usually occurs in lupus patients.

**Figure 7. Genes associated with risk for lupus in the apoptosis pathway.** The * represents genes which have been identified as having variants which increase risk for lupus. TNFα, CASP10, and IRF5 are pro-apoptotic whereas OPN and p21 are anti-apoptotic. These genes all have a role in how much apoptosis is occurring inside the cell. Once apoptosis has transpired, the cell must be cleared. Complement components such as C1q, C2, and C3b can opsonize dying cells or immune complexes for removal. Dying cells express altered lipids on their surface for recognition by proteins such as MFG-E8 and CRP. This system of recognizing apoptotic cells and immune complexes facilitates their removal by neighboring cells or immune cells.

The problem of creating autoantibodies could stem from too much apoptosis or too little clearance of apoptotic debris. Genes identified in GWA studies that could alter the amount of apoptosis in a cell include TNFα, caspase 10, IRF5, osteopontin, and p21.

TNFα was identified as a risk factor for lupus in certain ethnicities [119-120]. TNFα is a cytokine which is produced and secreted to signal to other cells and is found at high levels in the serum of lupus patients [121-123]. Part of its function is to induce apoptosis—when a cell binds TNFα, it activates the caspase cascade. Caspases are proteases which are activated in a cascade under certain conditions and are a hallmark of apoptosis. They cleave other caspases as well, and
the combined proteolytic activity of several different activated caspases breaks down cellular components as the cell prepares to die. Caspase 8 is activated by TNF signaling and cleaves caspase 10, which then cleaves caspases 3 and 7. Caspase 10 is another lupus susceptibility gene [95]. IRF5, as well as being a transcription factor which helps produce IFN, is also a tumor suppressor gene which has been shown to be inactivated in certain cancers [124-125]. This is because of IRF5’s pro-apoptotic function [126].

Osteopontin (OPN) and p21 are also lupus risk genes, both anti-apoptotic. OPN promotes proliferation, as well as prevention of death under apoptotic stimuli [127]. p21 is a cell cycle inhibitor that is normally activated in response to DNA damage. A mimic of p21 was used in the treatment of murine lupus in the NZB/NZW mouse, and it was found to dramatically reduce the disease [128].

Therefore, there are genes which dysregulate the amount of apoptosis, and they are associated with risk for lupus. But this is only half of the picture; the other part is the clearance of apoptotic cells or immune complexes (Figure 7). Several SLE susceptibility genes in this pathway have been identified as well. Active SLE can be assessed when low levels of complement proteins are found in circulation. Complement can function against microbes during an infection, but also help to degrade immune complexes—they can recognize and tag them for removal. Other non-complement proteins function to bind apoptotic cells or immune complexes to facilitate their uptake by other cells. This tagging for removal is called opsonization.
Integrin αM (ITGAM) has been convincingly associated to SLE [129]. Risk variants of ITGAM have been associated with certain clinical manifestations of lupus [130]. It is a cell receptor which binds to OPN or to complement C3b. C3b binds to apoptotic cells or immune complexes.

SLE association with polymorphisms in complement components C1q, C2, C4a, and C4b have large OR values, meaning that the risk alleles of these genes are causing a large effect. When C1q is expressed at low levels it can lead to lupus, and it was shown to increase the amount the of IFN produced due to immune complexes [131]. Complement components function by binding immune complexes by the Fc region of antibody or by binding to other parts of apoptotic cells, which can opsonize them for easier uptake by other cells. Cells can then remove the immune complex or apoptotic debris by endocytosis. Receptors for the Fc region of antibody have also been implicated in SLE risk [83]. These receptors can bind to antibody within an immune complex.

Other proteins such as milk fat globule epidermal growth factor 8 (MFG-E8) and C-reactive protein (CRP) can bind to apoptotic cells by recognizing phospholipids on their membranes. MFG-E8 binds to phosphatidylinerine, an “eat me” signal which is present on apoptotic cells. The MFG-E8 knockout mouse develops SLE because of failure to remove apoptotic cells [132]. CRP binds to phosphocholine, which is present on dying or damaged cells. Both MFG-E8 and CRP are lupus risk genes [133-135]. Low levels of mannose-binding lectin (MBL) can lead to higher levels of apoptosis in the case of this lupus-risk associated gene [136].
The number of genes associated with risk for SLE will likely increase, though we have an interesting pool of genes already that point to specific pathways associated with the disease. The interferon and apoptosis pathways are certainly important players in the etiopathogenesis of lupus.

1.10. Interferon regulatory factor 5

The gene of interest in this dissertation is human interferon regulatory factor 5 (Figure 8). IRF5 was discovered in 1998 by Paula Pitha [137] and first characterized by Betsey Barnes [138]. IRF5 is a transcription factor which binds to a GAAAN\(_N\)GAAA repeat, similar to other IRFs, which all contain a tryptophan repeat in the DNA binding domain [139]. IRF5 is expressed in B cells, macrophages, monocytes, and myeloid cells [21]. IRF5 is conserved among most vertebrates; including fish, fowl, amphibians, reptiles, and mammals [140].

Phosphorylation and dimerization are essential to activate IRF5 in its role as a transcription factor, and ubiquitination can also activate it [141-142]. Dimerization unmasks the nuclear localization signals (NLS) and masks the nuclear export signal (NES). There are two NLS and one NES sequences [143-145]. IRF5 can dimerize with IRF3 and IRF7, which have some similar and some distinct transcriptional targets [143, 146]. This dissertation focuses on interferon production and apoptosis, both of which are directly affected by IRF5. IRF5 is also involved in the production of other cytokines [147], B cell development [148], macrophage differentiation [149], and is a cell cycle inhibitor [150]. These important roles echo the many autoimmune diseases with which polymorphisms in the \textit{IRF5} gene are associated.
Figure 8. IRF5 mRNA and protein structures. (A) mRNA folding of IRF5 variant 2 (NM_032643.3) is shown and was performed using mfold [151]. The 5’ end of the mRNA is indicated, and a small hairpin forms at this end (see Section 2.4.5). (B) The protein folding shows a dimer of the crystal structure of the transactivation domain of pseudophosphorylated human IRF5 (PDB file 3DSH). The two colors represent the two monomers, and helix 5 extends from one monomer across to the other. This helix 5 is autoinhibitory to activation until the protein is phosphorylated and undergoes a conformational change. IRF5 can homodimerize or heterodimerize with other IRFs [152].

1.10.1. Autoimmune diseases associated with IRF5’s rs2004640 SNP and CGGGG indel

Linkage or association studies which compare disease groups and unaffected controls have revealed several loci in IRF5 that are associated with autoimmune disease. Of IRF5’s many disease-associated polymorphisms, only four have been identified as functional polymorphisms [112]. Two of these functional polymorphisms are in the promoter region of IRF5, and thus may directly affect IRF5 expression. This study examines these two functional promoter polymorphisms: the rs2004640 single nucleotide polymorphism (SNP) and the CGGGG indel—a copy number variant. There are other SNPs and indels within IRF5 that are associated to disease which will not be discussed, although several of the prominent ones are listed in Section 1.9.1.

The single nucleotide polymorphism (SNP) designated rs2004640 in IRF5 has been convincingly associated with systemic lupus erythematosus (SLE) in multiple ethnic groups [84, 90, 92, 106, 126].
More recently, this polymorphism has been associated with several other autoimmune diseases, such as rheumatoid arthritis [154-155], systemic sclerosis [156], multiple sclerosis [157], ulcerative colitis [158], and Sjögren’s syndrome [159].

The IRF5 risk T allele at rs2004640 is associated with altered symptoms in autoimmune diseases. For example, multiple sclerosis patients with the risk SNP have poor response to interferon-β therapy [160]. In rheumatoid arthritis, the rs2004640 T polymorphism is associated with anti-citrullinated protein antibody positivity [161]. The IRF5 risk polymorphism is associated with anti-topoisomerase antibody positive (ATA+) systemic sclerosis and lung disease [156, 162]. In lupus, the rs2004640 risk allele in IRF5 is associated with risk for nephritis [163] and the presence of dsDNA antibodies [164]. Importantly, risk haplotypes that include rs2004640 are known to correlate with higher IRF5 levels [112, 165], and correlate with higher cytokine activity, such as IFNα and TNFα [111, 166].

The CGGGG indel is associated with several autoimmune diseases, and part of the risk for either of the two polymorphisms could be due to their high linkage disequilibrium, which is around 0.55 ($r^2$) [85]. The CGGGG polymorphism is associated with several autoimmune or inflammatory conditions. It has been associated with SLE [84], Sjögren’s syndrome [108], multiple sclerosis [157], Crohn’s disease, and ulcerative colitis [158]. Inflammatory conditions may also be affected by the CGGGG indel. One study found it was associated with acute coronary syndrome [167].
1.10.2. The rs2004640 SNP and CGGGG indel

There are several kinds of splicing, including exon skipping, alternative exon usage, and others. Most kinds of alternative splicing do not affect transcription levels. However, in the case of alternative promoter splicing—using different first exons—there is a direct effect on the levels of transcription and translation. Each of the four exon 1 options (1A, 1B, 1C, and 1D) has a distinct transcriptional start site, and each is under the control of a different promoter. The rs2004640 and CGGGG risk polymorphisms directly affect the levels of exons 1A and 1B of IRF5 due to their position in the promoter region (Figure 9).

![Figure 9. IRF5 mRNA and the position of the rs2004640 SNP and CGGGG indel.](image-url)

**Figure 9. IRF5 mRNA and the position of the rs2004640 SNP and CGGGG indel.** The genomic region is shown with exons as boxes and introns as lines. The protein coding and untranslated regions are shown at the top. rs2004640 is at the splice acceptor site for exon 1B, and the CGGGG indel is 64bp upstream from the transcription start site for exon 1A. Drawn to scale, but with introns reduced in size 10:1. UTR: untranslated region, SNP: single nucleotide polymorphism

The rs2004640 SNP is a G/T polymorphism, where the T risk allele creates an alternate splicing site for exon 1 [153]. This change creates a sequence-specific recognition site for the spliceosome, allowing use of an alternate first exon, called exon 1B (Figure 10A). The T nucleotide is within the intron between exons 1 and 2, and it is removed from the mRNA upon splicing. Without the risk T allele, exon 1B cannot be spliced onto exon 2 and would encode a
null-functional protein. These transcripts are usually targeted by non-sense mediated decay [168].

The CGGGG indel is located within the promoter for exon 1A, 64bp upstream of its transcription start site (Figure 10B). There are either three (3X) or four copies (4X) of the CGGGG repeat sequence; four copies is the risk variant. This additional copy allows binding of additional transcription factors [85]. The CGGGG 4X variant is associated with higher levels of the cytokines TNFα, IL-12p40, IL-8, IL-1b, and IL-10 [166]. IRF5 levels may be affected by this indel. One study in thymic tissue found no association between the CGGGG indel and IRF5 levels, although there was decreased usage of exon 1A [169]. Another study of peripheral blood mononuclear cells (PBMCs) showed increased IRF5 levels from CGGGG 4X risk cells [85]. The CGGGG indel is also designated rs77571059. It has been described as a GGGGC repeat, since C’s flank the repeat sequences.

IRF5 exons 1A, 1B, 1C, and 1D each have a distinct transcriptional start site, and as is the case with every first exon, each exon 1 of IRF5 has its own promoter. The promoter for each exon is not well studied, although previous work has shown that promoters for exons 1A and 1C are controlled by an IRF element (IRFE) and an interferon stimulatory response element (ISRE), respectively [170]. Although most alternative splicing does not directly affect transcription levels, alternative promoter splicing—using different first exons—does directly affect the levels of transcription. IRF5’s first exon is part of the 5’ untranslated region (5’UTR) and thus does not affect the amino acid sequence. Also, using different first exons has not been shown to contribute to alterations in amino acid sequence of the IRF5 protein [90]. Instead, the effect of including
using different promoters and first exons is likely to be on the mRNA—such as the level of transcription, alternative splicing, the mRNA secondary structure, the stability of the mRNA, and the efficiency of its translation. All of these effects were investigated in the present study. The

Figure 10. Functional changes due to IRF5 promoter polymorphisms. (A) The position of the rs2004640 SNP on pre-mRNA. Before splicing, the messenger RNA has either a U (encoded by the risk T allele) or a G. The colored letters shown at the bottom of panel A are a WebLogo, which represent the consensus recognition sites for the spliceosome. The height of the stack represents how often those nucleotides are found at that position, and thus the high GT represents a strong preference for recognizing GT at the intron boundary. This matches in the risk T allele (GT at the intron boundary), but not the protective allele (GG at the intron boundary). A person homozygous for the protective allele cannot splice IRF5 mRNA that begins with exon 1B. Instead of a functional protein, the resultant mRNA would encode a junk protein and be targeted for non-sense mediated decay. Splice junction WebLogos are from Stephens, et al. [171]. (B) The CGGGG indel is an insertion/deletion of a CGGGG repeat upstream of exon 1A, and it is part of exon 1A’s promoter. When there are four copies, additional SP1 transcription factors—which bind to GGCGG—can bind to the promoter, altering transcription levels.
mechanisms of mRNA changes will better explain the pathways that are misfiring in SLE and other associated autoimmune diseases.

1.11. Conclusions to the introduction

SLE is a prototypic disease for studying autoimmunology. Apoptosis, clearance of apoptotic cells, the formation of immune complexes, and the production of interferon are all tied to SLE as either causes, effects, or both. The actual genetic causes of autoimmune diseases such as SLE are being identified through genome wide association studies, and then individually examined. This strategy will lead to better understanding and thus better treatments for these complex diseases.

The gene IRF5 contains several polymorphisms associated with autoimmune diseases including SLE. Of IRF5’s many disease-associated polymorphisms, only four have been identified as functional polymorphisms. Two of these functional polymorphisms are in the promoter region of IRF5, and thus may directly affect IRF5 expression. The two functional promoter polymorphisms, the rs2004640 SNP and the CGGGG indel may exhibit functional changes which will teach us more about the pathways altered by them. The data presented in summary above (Section 1), as well as the research summary and discussion below (Sections 2 and 3), support the hypothesis that changes in IRF5 signaling promote an autoimmune state in those with the genetic propensity.
2. Research sections

2.1. *IRF5* genotyping and volunteer demographics

Convenience sampling was performed under IRB approval to collect blood samples from volunteers on or around Brigham Young University (BYU) campus. Demographic data were collected to determine gender and ethnicity in order to pair samples. Matching people by other variables and only varying the rs2004640 genotype allows the creation of matched sets and the use of the Wilcoxon signed rank sum test, which is appropriate for paired data which is non-parametric [172].

For the use of the samples, three separate items were created: serum, DNA, and a lymphoblastoid cell line (LCL). The sample of serum was collected from coagulated blood after centrifugation (although no data presented herein were obtained by tests of serum). A sample of genomic DNA was extracted from peripheral blood mononuclear cells (PBMC). These PBMCs, also known as a buffy coat, were removed from blood by ficoll density gradient. Finally, an immortalized cell line of LCLS was established for each volunteer by infecting the PBMCs with Epstein-Barr virus in the presence of cyclosporin in order to kill T cells. At least two weeks were allowed to establish the cell lines before testing began.

These immortalized cells generated from PBMCs grow indefinitely in culture and allow for a repeatable method of testing. Although immune signaling may be dysregulated in LCLs, which are transformed by EBV, both risk and protective cells lines always received the same treatment when testing and when maintaining cells. LCLs have been found to have very low (0.3%)
somatic mutation rate [173] and a 96% concordance rate of SNP calls before and after transformation [174]. They are an excellent and convenient model system.

Genotyping was performed by TaqMan-based real-time PCR of genomic DNA with a two-color probe set for verifying probe-quencher liberation of either or both colors. Sets of two people were created by finding a homozygous risk T/T sample and homozygous protective G/G sample and matching on gender and ethnicity. Heterozygotes were not included in this study. Genotyping at rs10488631 was also performed using a similar TaqMan-based system, but the rs10488631 genotype was not used in forming matched sets; rather it was used for analysis.

Genotyping at the CGGGG indel was performed using primers which flank the indel for touchdown PCR, followed by size analysis on a 4% agarose gel. The limitation of this method is the apparent preference for the protective genotype’s PCR product. The protective product seems to predominate, such that no heterozygous individuals can be reasonably determined, either by size determination or sequencing. CGGGG genotyping was only performed on those already in rs2004640 sets.

The genotyping results from the three loci are shown in Figure 11. The rs2004640 risk allele is found in many individuals (80%), and 30% had the risk factor on both alleles. The CGGGG risk 4X variant was only found in those who had the rs2004640 T/T allele, and in about half of those with the protective G/G allele at rs2004640. There was a 55% linkage disequilibrium of CGGGG to rs2004640 in a previous study [85], and our concordance ($r^2$) was 50%.
Figure 11. Demographics of samples collected for testing. For matching rs2004640 risk T/T to rs2004640 protective G/G, matches were the same gender and ethnicity. The other two genotypes were used for stratifying during statistical analysis. Healthy individuals and several lupus patients were genotyped, although no lupus patients were used for testing since no lupus patients were found with the protective allele at rs2004640. (A) Using a TaqMan-based genotyping assay, the prevalence of the risk T/T allele in rs200460 was 30%, with 20% protective G/G. These homozygotes were further subtyped by the CGGGG indel found in the 1A promoter (see Figure 11E), although the risk variant which has four copies of the indel was only found among those with the risk rs2004640 allele (42%). (B) Using a TaqMan-based genotyping assay, the prevalence of the rs10488631 risk C/C allele was 16%, and the protective T/T allele 67%. (C) Of all the samples collected most were female (86%). (D) Ethnicities are shown as well. Most were European American (64%) (E) Genotyping results at the CGGGG indel. Individuals have either three or four copies of a CGGG repeat. Genotyping was done by amplification using touchdown PCR, and size determination on a 4% agarose gel. The primers should yield either a 103 or 108bp band. Numbers are the identifier for each cell line. L: Ladder, 100bp band shown.

Once matched sets were established to compare risk and protective individuals, tests were performed to evaluate the effect of the risk factor on apoptosis, splicing, mRNA stability, and other gene’s expression levels.
2.2. *IRF5*’s four promoters

IRF5 uses one of four first exons for each molecule of mRNA—either exon 1A, 1B, 1C, or 1D. The four first exons of IRF5 would be actively transcribed or not depending on the cellular transcription factors that are able to bind it. An analysis of the promoters for each of the four first exons of IRF5 was performed. First a list transcription factors that have been shown to actually bind the promoter was generated by using the encyclopedia of DNA elements chromatin immunoprecipitation sequencing (ENCODE ChIP-Seq) data set [175]. This list represents many experiments’ results of transcription factors that bind to this genomic region of DNA (Figure 12A). The transcription factors listed have also been associated with specific binding sites. WebLogos, which visualize consensus binding sites [176], were generated for each ENCODE transcription factor tested and compiled in the online database FactorBook [177]. The consensus sites were converted manually into an ambiguous DNA code search term, where for example W (weak) represents an A or a T nucleotide (Figure 12B). The consensus search term was then used to search the proximal promoters (~200bp upstream from the +1 sites) to encounter a proposed binding site. Consensus search term screening was performed using MEGA [178]. Several transcription factors binding sites were found in the regions upstream of transcription start sites (Figure 12C). The start sites were taken from reference sequences for exons 1A, 1B, and 1C, and the sequence for variant 12 of IRF5 for exon 1D (no reference sequence exists at present for exon 1D). The source sequences are GenBank IDs: NM_002200.3, NM_032643.3, NM_001098627.2, and EU258897.1 for exons 1A, 1B, 1C, and 1D, respectively. The workflow and results are shown in figure 12, with TCF12 as an example WebLogo.
Figure 12. Promoter analysis results for each first exon of IRF5. (A) ENCODE data shows ChIP-Seq analysis of the promoter region of IRF5. This data was used in conjunction with the FactorBook database to define nucleotide searches on the genomic sequence. (B) The consensus search terms generated from FactorBook, with the TCF12 consensus binding site as an example. (C) The final analysis of potential binding sites is shown along the genomic DNA promoter sequences, with color-coded boxes representing the binding sites or transcription factors shown in the key. AP: activator protein, BRE: B-response element, CTCF: CCCTC binding factor, TCF: transcription factor, EBF: early B cell factor, IRF: interferon regulatory factor, NFκB: nuclear factor kappa light chain enhancer of activated B cells, PAX: paired box, PU: purine rich, SP: specificity protein, STAT: signal transducer and activator of proteins, TATA: thymidine adenine
The promoters for the four first exons of *IRF5* have different potential binding sites. The 1A promoter contains putative binding sites for PAX5, PU.1, SP1, and TCF12 which binds to enhancer boxes (E box). An extra SP1 binding site appears in those with the CGGGG 4X indel. Exon 1B’s promoter was the only *IRF5* promoter with a p53 binding site. This is discussed in more detail in Section 2.2.3 below. 1B also has SP1, TCF12, IRF4, and EBF sites. The 1C promoter was the only promoter with STAT2, AP1, and Myc binding sites; it also has SP1 and IRF4 sites. The 1D promoter evaluation showed only four potential binding sites for transcription factors: SP1, CTCF, IRF4, and NFκB.

The promoters were cloned using PCR to transfer onto a luciferase reporter plasmid. There are two versions of the 1A promoter, one having the 4X variant of the CGGGG indel (1A risk) and the other having the 3X variant (1A protective). The 1B promoter was cloned using a nested PCR. The first round primers annealed outside of an inverted repeat mentioned in Section 2.2.2 below, and the inner nested primers amplified the promoter itself. The 1D promoter was successfully cloned only after the current writing, therefore it was not included in the current analysis.

A luciferase assay was performed using the pGL4 plasmid, with the promoters of *IRF5* placed in front of luciferase. The activity levels of the promoters were analyzed in different cell types since different transcription factors would be active in different cells. Immune cells were electroporated, or epithelial cells were calcium phosphate cotransfected with one of the luciferase plasmids as well as a plasmid which expresses a green fluorescent protein (GFP),
which was included as a transfection control. Values for luciferase expression were then normalized to account for transfection efficiency.

The activity level of four promoters was measured in HEK293T cells, a human embryonic kidney line expressing the Simian virus 40 large T antigen. This represents potential expression levels of each promoter in epithelial cells. Raji cells are a Burkitt’s lymphoma B cell line and represent potential expression in immune cells. The 1A risk promoter showed high activity in both cell types. The promoter for exon 1B showed moderate activity, and the 1C promoter was

---

**Figure 13. Transcriptional activity of each promoter of IRF5.** Preliminary data from one experiment shows activity of the promoters listed on the x-axis. The luciferase plasmids were cotransfected by the calcium phosphate transfection method into HEK293T cells and by electroporation into Raji cells. A control green fluorescent protein-expressing plasmid was measured (relative fluorescence units, RFU) in each sample to normalize transfection efficiency. The levels of transcription were lower in epithelial HEK cells, where IRF5 is not normally expressed [142], and higher in Raji cells, a B cell line (IRF5 is normally expressed in B cells [21]). HEK: human embryonic kidney, RLU: relative luminescence units.
transcribed at the lowest level in both cases (Figure 13). This data should be considered preliminary, since it was not repeated.

2.2.1. Activity of IRF5’s exon 1A and 1D promoters affected by TLR7 ligation

IRF5 expression increases due to several signaling pathways including TLR7. TLR7 is an endosomal receptor, as opposed to the majority of TLRs which are expressed at the cell surface. Endosomal TLRs such as TLR7 require the ligand to be first endocytosed into the cells, and then merged with the endosome that contains TLR7. Most endosomal TLRs bind to nucleic acids. Single stranded RNA is the natural agonist for TLR7. TLR7 can also be activated by the small synthetic compounds such as the imidazoquinolines called imiquimod and resiquimod. Imiquimod is a TLR7 ligand and resiquimod is a ligand for TLR7 and TLR8 [179]. Imiquimod is used clinically as a topical cream for treatment of genital warts and certain cancers. It activates the immune system, which recruits inflammatory mediators to kill the virus-infected or cancerous cells [180]. Imiquimod treatments were thus performed to increase IRF5 levels, and then determine which promoter was being activated.

Cells were treated with imiquimod at 25μg/ml for 24h, and then cDNA was prepared from an RNA extract. This was done for lymphoblastoid cell lines (LCLs) generated from twenty healthy individuals. As expected, the levels of IRF5 increased by 1.9 fold when cells are treated with imiquimod (Figure 14A).

The amounts of each first exon were measured and normalized to both a housekeeping gene and to the total amount of IRF5. This measures both the level of each first exon and the proportional
**Figure 14. Activation of IRF5 transcription through exon 1A and 1D promoters upon imiquimod treatment.** All mRNA levels were measured in LCLs generated from healthy individuals. Since 1B primers didn’t work in the SYBR green-based assay, the levels of IRF5 first exons were measured in both a SYBR green- and TaqMan-based quantitative PCR. (A) IRF5 levels were measured using SYBR green-based quantitative PCR. The levels of IRF5 were 1.9-fold higher in treated cells (p=0.0002) after normalizing to a housekeeping gene. (B) Using TaqMan-based quantitative PCR, the levels of exons 1A increased 2.2 fold (p=0.030) and 1D increased by 2.8 fold (p=0.033) after normalizing to a housekeeping gene. (C) Using SYBR green-based quantitative PCR, the levels of 1C were lower by 2.3-fold (p=0.0064) after normalizing to total IRF5 levels. (D) Using a TaqMan-based quantitative PCR assay and normalizing to total IRF5 levels, the levels of 1B and 1C were decreased, however these data did not reach statistical significance. Taken together, the levels of exons 1A and 1D increase, and in turn the levels of 1B and 1C are proportionally lower. Numbers in parentheses are the sample size. Statistical significance determined by the Wilcoxon signed rank sum test. IRF: interferon regulatory factor.
level of each first exon. By including a comparison of the total to the sum of its parts, we get a better picture of proportional changes in first exon usage.

The levels of exons 1A and 1D increased by at least two fold when compared to a housekeeping gene (Figure 14B). When the levels of each first exon were compared to the total IRF5 levels (and not to a housekeeping gene), exon 1B and 1C levels decreased and the others remained the same (Figures 14C and 14D). Taken together, the increase in total IRF5 is due to increased overall usage of 1A and 1D.

2.2.2. Inverted repeat encompasses IRF5 exon 1B and its promoter

During cloning experiments dealing with exon 1B and its promoter, several sequencing reactions showed less than 100% sequence identity to the target. It was soon discovered that the primers were annealing to an upstream inverted repeat sequence. This repeat necessitated nested PCR for cloning the 1B promoter and ordering synthesized oligos instead of cloning the 1B 5’UTR (Table 4). The repeat length is 1.8kbp, and the two copies have 82.8% identity [181] (Figure 15). The function of this repeat is unknown, but repeated sequences can act as decoys for transcription factors, lowering transcription of the intended target [182].

![Figure 15. Inverted repeat in the promoter region of IRF5.](image)

The repeat is 1.8kbp long, and the two versions match at 82.8% of bases. The right repeat begins in the last 80bp of exon 1A. It extends through the promoter for exon 1B, exon 1B itself, and then 1473bp past exon 1B. Exons are shown in blue. Drawn to scale.
2.2.3. Putative p53 binding site in IRF5’s exon 1B promoter

The promoter analysis described in Section 2.2 above revealed a potential p53 binding site. p53 binds as a tetramer to two copies of the sequence rrrCwwGyyy, with a spacer of 0-13 nucleotides between the copies [183]. A near match for this sequence was found in the 1B promoter (Figure 16B). This is very important due to the role of apoptosis dysregulation in SLE. p53 is a main controller of apoptosis, which is why it is so commonly mutated in cancers [184]. IRF5 is also pro-apoptotic in a p53 independent manner, and thus if p53 activates IRF5, apoptosis levels will be additively altered. If p53 can actually control the 1B promoter, and since the 1B promoter is only used in cells with the rs2004640 risk allele, apoptosis would be altered in rs2004640 risk cells. This could be a key to understanding the source of autoimmune risk when using exon 1B.

In all body cells p53 is constantly transcribed and translated, but also constantly sequestered and degraded. Signals of DNA damage, oxidative stress, or abnormal cell cycle can all activate p53. p53 is activated by removing the inactivating protein MDM2. p53 is further activated by phosphorylation, tetramerization, and translocation to the nucleus. It then acts as a transcription factor to alter gene expression. Its three-fold mission is to activate DNA repair, halt the cell cycle, and activate apoptosis. In this seemingly contradictory set of activation programs, the cell is induced to either repair and live or fail to repair and die [185].

To test p53 activity in the 1B promoter compared to the other promoters, HEK293T cells were calcium phosphate transfected with the promoter luciferase plasmids. After allowing 24 hours for expression, DNA damage was induced by treating the cells with UV light (75 cm from a 254 nm UV bulb for 90 seconds) and incubating the cells for three hours; as a control some cells received
no UV exposure. The amounts of luciferase were then measured to determine promoter activity. The luciferase levels decreased for the 1A promoters and the 1C promoter, but increased for the 1B promoter (Figure 16A). The decrease may be due to the cell cycle-inhibitory and apoptosis-inducing signaling induced by UV damage. These data should be considered preliminary, since a control for transfection efficiency was not included in the analysis.

Figure 16. p53-dependent activity of IRF5’s 1B promoter upon DNA damage. (A) Preliminary data from three replicates of a luciferase reporter assay. HEK293T cells were transfected with the indicated promoter plasmids. Promoters for exons 1A and 1C decrease in activity three hours after a 90 second UV exposure, but the 1B promoter of IRF5 increases in activity. (B) The putative p53 binding site in IRF5’s exon 1B promoter, with a WebLogo of the p53 consensus binding site [186] to indicate important bases and matches. The height of the base represents the frequency of that nucleotide. Site-directed mutagenesis was performed to mutate the binding site at the critical C and G bases as highlighted by boxes. (C) Preliminary data from a luciferase reporter assay. Three different LCLs generated from healthy volunteers were electroporated with the 1Bwt or the 1Bp53* promoter luciferase plasmid. Cells were treated with 0.5 mM etoposide or left untreated. The plasmid with the intact (wild type) p53 binding site shows an increase in activity, but the mutated version that should not be able to bind p53 decreases after a 48 hr etoposide treatment. UV: ultraviolet, wt: wild type, *: mutant
To test whether the p53 binding site is the cause of the increase, the 1B plasmid was mutated using site-directed mutagenesis to change the wild type sequence to an altered sequence to which p53 should not be able to bind. The consensus binding site contains four conserved C or G bases which were mutated to A or T on the luciferase plasmid (Figure 16B). The wild type and p53-mutant luciferase plasmids were transfected by electroporation into three different LCLs generated from healthy volunteers. After 24 hours to allow for plasmid expression, cells were either treated with etoposide or left untreated for 48 hours. The levels of luciferase activity again increased with the 1B promoter, but decreased when using the p53 mutant version. Again these data should be considered preliminary, since a control for transfection efficiency was not included in the analysis.

The p53 binding site appears to be able to bind p53. Therefore we sought to test apoptosis levels in cells with the rs2004640 risk T/T genotype compared to the protective G/G genotype.

2.3. The rs2004640 SNP’s role in apoptosis

As described in Section 1.10.2 above the rs2004640 risk T allele allows use of exon 1B. Furthermore, the promoter for exon 1B contains a p53 binding site (see Section 2.2.3). Taken together, there should be altered levels of apoptosis in cells that can use exon 1B—those with the rs2004640 risk T allele. To this end, apoptosis was induced, and counts of apoptotic cells were measured in the matched sets described above (Section 2.1).

Four apoptotic agents were used to increase apoptotic death in LCLs. They were etoposide, 5-fluorouracil (5FU), activating antibodies to TNFα-related apoptosis-inducing ligand (TRAIL),
and activating antibodies to Fas receptor. Etoposide and 5FU both induce intrinsic apoptosis; the TRAIL-Ab and Fas-Ab both induce extrinsic apoptosis. Etoposide is a topoisomerase inhibitor which causes dsDNA breaks [187]. 5FU is a nucleotide analog which blocks thymidine synthesis [188]; TRAIL-Ab and Fas-Ab are activating antibodies which mimic ligation of these death receptors. As a control for equalizing the drug effect on individual cells in the population, hydroxyurea (HU) was used to synchronize the cell cycle in the cell population. Cells were treated for 40 hours with 0.2 mM HU, which was washed off before treatment. All the data collected were the apoptotic index, which is the change in apoptosis levels compared to untreated cells.

For detection of apoptosis, cells were assayed using a dual-color flow cytometry analysis. Cells that were positive for Annexin V, which binds to phosphatidyl serine (PS), were considered to be in early stages of apoptosis, since PS flipping to the outer leaflet of the cell membrane is an early apoptotic event [189]. Annexin V itself is not fluorescent, so a fluorescein isothiocyanate (FITC) conjugate is attached to it. Cells positive for propidium iodide (PI) are considered late-stage apoptosis, since this DNA intercalator is not membrane permeable, and therefore only dead cells without an intact membrane will stain [190].

The levels of apoptosis were lower in rs2004640 T/T risk cells when a low concentration of etoposide was used. However, when an increased dosage of etoposide was used, the levels of apoptosis were higher in cells with the risk allele. Large doses of etoposide yielded more apoptosis in risk cells after two days (Figure 17); this is consistent with the hypothesis that p53 activation can increase apoptosis. However, apoptosis levels were higher in protective cells when
small doses of etoposide were administered for two days (Figure 17), and after 12 hour 5FU treatments (Figure 18). In either case, the risk allele did alter the levels of apoptosis in LCLs after inducing DNA damage.

Figure 17. Altered apoptosis levels due to the rs2004640 polymorphism in IRF5. LCLs from healthy individuals were cell cycle-synchronized by treating with 0.2 mM hydroxyurea for 40 hours. This was removed by washing the cells before applying the etoposide at the concentrations indicated for 48 hours. To detect apoptosis, propidium iodide (PI) and Annexin V were applied. Propidium iodide stains DNA in cells without an intact cell membrane, and Annexin V binds to phosphatidylserine (PS) present on the outer leaflet of the cell membrane. Cells that are healthy maintain the PS on the inside on the cell on the inner leaflet of the membrane and remain unstained. Annexin V was conjugated to the fluorescent molecule fluorescein isothiocyanate (FITC). (A) Compared to cells with the protective G/G allele, cells with the risk T/T allele at rs2004640 exhibited a decrease in apoptotic cells at 0.1 mM etoposide, but an increase at 1 mM etoposide, as indicated by the levels of Annexin V-FITC fluorescence or PI fluorescence. (B) A representative plot from one matched set is shown from each set of experiments. Statistical significance determined by the Wilcoxon signed rank sum test.

The apoptosis pathways affected by the IRF5 rs2004640 SNP appear to be the intrinsic pathways only. Cells were treated with either the intrinsic apoptosis inducer 5FU or by activating extrinsic apoptosis through the death receptors TRAIL and Fas. When comparing rs2004640 risk to
protective LCLs, no difference in apoptosis was found for extrinsic apoptosis inducers, but the levels of apoptosis were lower in risk cells treated with 5FU (Figure 18).

<table>
<thead>
<tr>
<th>5-Fluorouracil (1.5mg/ml)</th>
<th>Anti-TRAIL Ab (1ug/ml)</th>
<th>Anti-Fas Ab (5ug/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2004640 Risk (T/T) N=8</td>
<td>rs2004640 Risk (T/T) N=8</td>
<td>rs2004640 Risk (T/T) N=8</td>
</tr>
<tr>
<td>Protective (G/G) N=8</td>
<td>Protective (G/G) N=8</td>
<td>Protective (G/G) N=8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p=0.036</td>
<td>p=0.33</td>
<td>p=0.56</td>
</tr>
</tbody>
</table>

**Figure 18. Altered intrinsic but not extrinsic apoptosis levels in rs2004640 risk T/T cells.** LCLs from healthy individuals were cell cycle-synchronized by treating with 0.2 mM hydroxyurea for 40 hours. This was removed by washing the cells before applying the treatments at the concentrations indicated for 12 hours. To detect apoptosis, propidium iodide (PI) and Annexin V were applied. Propidium iodide stains DNA in cells without an intact cell membrane, and Annexin V binds to phosphatidylserine (PS) present on the outer leaflet of the cell membrane. Cells that are healthy maintain the PS on the inside on the cell on the inner leaflet of the membrane and remain unstained. Annexin V was conjugated to the fluorescent molecule fluorescein isothiocyanate (FITC). Compared to cells with the protective G/G allele, cells with the risk T/T allele at rs2004640 exhibited a decrease in apoptotic cells at 1.5 mg/ml of 5-fluorouracil, but no change at 1 μg/ml TRAIL-activating antibody or at 5 μg/ml Fas-activating antibody, as indicated by the levels of Annexin V-FITC fluorescence after the 12 hour treatment. Statistical significance determined by the Wilcoxon signed rank sum test. TRAIL: tumor necrosis factor-related apoptosis-inducing ligand

### 2.4. Effect of risk polymorphisms on IRF5 mRNA

The rs2004640 SNP and the CGGGG indel are in the promoter region of *IRF5*. Their position on the genome, however, does not necessarily cause any change to the amino acid composition of
the IRF5 protein. Any changes will be in protein levels, mRNA levels, or mRNA composition. Tests of this hypothesis are in the following sections (Sections 2.4.1 through 2.5).

2.4.1. Increased IRF5 expression levels in cells with the rs2004640 risk T/T allele and the risk 4X allele at the CGGGG indel

Both the CGGGG 4X variant and the rs2004640 T variant have been associated convincingly with autoimmune disease. Both have been dubbed functional polymorphisms, since a logical functional effect can be suggested for each. The proposed source of association with autoimmune disease is an increase in the levels of IRF5. The rs2004640 risk T polymorphism increases IRF5 levels in LCLs and PBMCs [112, 165], and the CGGGG risk 4X polymorphism has been shown to increase IRF5 in PBMCs [85], but decrease 1A-specific IRF5 transcripts in thymic cells [169]. We measured IRF5 transcription levels in order to compare average expression and perform regression modeling. This identifies the source of changes according to each polymorphism’s effect.

Quantitative PCR was performed on cDNA from cells with risk or protective alleles at each polymorphism. Healthy individuals were used, which removes potentially confounding factors due to existing autoimmune disease. The cells used were from volunteers that were all female, with 58% European American, 33% Hispanic, and 8% Asian. PBMCs were converted into lymphoblastoid cell lines (LCL) by transformation with Epstein-Barr virus (EBV). LCLs are transformed B cells, a cell in which IRF5 is normally expressed [21] and thus a good model.
mRNA expression levels of IRF5 were 2.7-fold higher in those with the rs2004640 risk allele (p=0.0018), when normalized to the levels of β-glucuronidase. For the CGGGG indel, risk cells were 2.1-fold higher (p=0.030) (Figure 19A). Western blots were performed on protein extracts from risk and protective cells. The levels of IRF5 protein were 1.8 fold higher in rs2004640 risk cells (p=0.040) and 1.6-fold higher in CGGGG risk cells (p=0.053, which only approaches significance), supporting the mRNA data (Figure 19B). For both the protein and mRNA data, regression modeling using a backward selection showed that the rs2004640 was the better...
predictor of IRF5 levels, with the better model excluding the CCGGGG effect. The CCGGGG indel was not sufficient alone to explain IRF5 mRNA or protein levels (ANOVA, p>0.05), although the rs2004640 SNP was sufficient alone (ANOVA, p<0.05). IRF5 levels were not affected by the rs10488631 genotype (data not shown), although our sample size for the risk allele at this locus was small.

### 2.4.2. IRF5 mRNA stability not affected by the rs2004640 SNP or CCGGGG indel

One further test of IRF5 mRNA was to evaluate its stability in cells. Actinomycin D was applied, which inhibits RNA polymerase and thus mRNA production. IRF5 mRNA levels were measured

![Figure 20. IRF5 mRNA stability not affected by the rs2004640 SNP or CCGGGG risk polymorphisms. (A) LCLs generated from healthy individuals were treated with Actinomycin D to measure IRF5 mRNA half-life comparing cells with the rs2004640 risk allele to those without. A SYBR green-based quantitative PCR was used to measure total IRF5 mRNA levels and normalized to GAPDH mRNA levels. No statistical difference was observed between the slopes comparing rs2004640 risk to protective. However, IRF5 mRNA expression in rs2004640 risk cells was consistently higher (p<0.001) by the Wilcoxon signed rank sum test. (B) Actinomycin D-treated LCLs were used to measure mRNA half-life of each first exon of IRF5. All cells included in the analysis had the risk T/T allele at rs2004640, so that the stability of exon 1B could also be measured (exon 1B is not used in cells with the G/G allele). A TaqMan-based quantitative PCR was used to measure first exon levels normalized to total IRF5. No statistical difference was observed among slopes by paired t-test. Values are changes in threshold cycle (ΔCₜ) which are high when expression is low. Therefore, for ease of viewing the y-axis is reversed and the x-axis was set below the highest value. All error bars represent standard error. Similar results were obtained for the CCGGGG indel (not shown).](image-url)
at two, six, and 24 hours after treatment of LCLs. Using a SYBR green-based quantitative PCR, the IRF5 mRNA levels were consistently higher in rs2004640 risk cells (p<0.001). However, no statistically significant difference in half-life was observed—all cell types had the same half-life slope over 24 hours (Figure 20A). Similar results were found for the CGGGG polymorphism as well (data not shown).

Using the same actinomycin D-treated samples, the relative stabilities of the first exons of IRF5 were evaluated using a TaqMan-based quantitative PCR assay by comparing the first exon to the total IRF5 expression level (Figure 20B). Cell lines that could use exon 1B (rs2004640 T/T) were evaluated so that the relative levels of exon 1B could be assessed with the other first exons. The mRNA was more difficult to detect when amplifying the first exon, with no first exons measurable by 24 hours. No first exon seemed to confer added stability compared to other first exons, since no statistically significant differences were observed among slopes.

2.4.3. Altered exon 1 usage in cells with the rs2004640 risk allele

The first exon of IRF5 can be any of four options: exons 1A, 1B, 1C, or 1D. This alternative promoter splicing may be used so that a single protein may be produced in response to many different stimuli. Depending on the cell type and the signals the cell receives, any or all of the promoters may be activated to produce IRF5. The expression level of each first exon was measured in unstimulated LCLs generated from healthy volunteers.
For a proportional analysis of first exon usage, levels of each first exon were measured by TaqMan-based quantitative PCR and normalized to the total amount of IRF5. To calculate proportions, the inverse of the ΔC_T was used since C_T values are low when expression is high.

**Figure 21.** Decreased usage of exon 1C and 1D and increased 1B in cells with the rs2004640 risk allele; decreased exon 1D usage in cells with the CGGGG risk allele. (A) TaqMan-based quantitative PCR measured IRF5 mRNA expression levels specific to first exon usage. The total amount of IRF5 mRNA was used to normalize the levels, thus showing a proportion of the total IRF5 expression. Since ΔC_T levels are low when expression is high, the inverse of the ΔC_T value was used to generate the graphs. The area of each graph is also proportional to the overall IRF5 expression levels (see Figure 19A). In lymphoblastoid cells lines (LCLs) generated from healthy individuals with the rs2004640 risk T/T allele, exon 1B was used in 22% of IRF5 mRNA, which is absent from cells with the protective G/G genotype. Numbers in parentheses are sample size. (B) SYBR green-based quantitative PCR measured IRF5 mRNA expression levels specific to first exon usage, in proportion to total IRF5 expression. Usage of exon 1C was 2.9-fold lower (p=0.026) and 1D was 2.8-fold lower (p=0.0056) in LCLs with the risk compared to the protective allele at rs2004640 by the Wilcoxon signed rank sum test. LCLs with the risk compared to the protective version of the CGGGG indel had 3.2-fold lower 1D levels (p=0.00055) by t-test. Values are calculated using the 2^ΔΔC_T method (fold difference) by setting the lowest expression for each genotype to 1. Error bars represent standard error.
Considering cells homozygous for the 4X CGGGG polymorphism (4X/4X), the proportions of the first exon usage did not differ greatly compared to 3X/3X cells, with exon 1D being having the greatest change, 5% lower in risk (Figure 21A). The effect of the rs2004640 genotype was more pronounced, due to exon 1B usage only in risk cells. In cells with the protective SNP, 35% of total IRF5 mRNA contained exon 1A, 28% contained exon 1C, and 37% had 1D. In cells with the risk allele 1A was found in 30% of mRNA, only 20% used 1C, and only 28% used 1D. Exon 1B was used exclusively in risk cells and comprised 22% of the 1/ΔC\textsubscript{T} levels of mRNA (Figure 21A).

Relative mRNA concentrations were compared between risk allele-containing and protective allele-containing cells using a SYBR green-based quantitative PCR assay. Total IRF5 was used to normalize the levels for a proportional analysis. For the rs2004640 SNP, 1C levels were 2.9-fold lower in risk (p=0.026) and 1D levels were 2.8-fold lower in risk (p=0.0056). The CGGGG indel affected only exon 1D, which was 3.2-fold lower (p=0.00055) (Figure 21B).

### 2.4.4. Effect on translation due to usage of the four distinct 5’UTRs of IRF5

Testing the effect of the rs2004640 and CGGGG risk polymorphisms moved downstream from the promoter to the mRNA. Since the rs2004640 risk allele affects which first exon of mRNA can be used and the CGGGG indel affects the levels of usage, we surmised that the changes would be measurable in mRNA studies. *IRF5* can encode the same protein regardless of the first exon with which it begins [90]. Assuming there is no effect on the protein—since the first exon is non-coding—we performed experiments to assay the effect of the 5’UTR, which can affect translation of the mRNA.

53
Since IRF5 mRNA may begin with one and only one of four different first exons, we evaluated the translational efficiency of mRNA depending on which first exon was used. It should be noted that exon 1B can only be used by individuals with the risk allele at rs2004640, but the SNP itself was not included on the luciferase plasmid (described in the next paragraph), since the SNP is eliminated from the RNA upon splicing onto exon 2.

Different from the promoter luciferase assays described above (Figures 13 and 16), 5’UTR luciferase assays were performed to determine the effect of different 5’UTRs on translation efficiency. The 5’UTRs of IRF5 contain exon 1A, 1B, 1C, or 1D and the untranslated part of exon 2. Each was inserted into a luciferase vector and used to compare the effect on translational efficiency. Instead of different promoters, the luciferase expression was driven by the SV40 promoter in all cases. HEK293T cells, which do not express IRF5 normally [142], were transfected by the calcium phosphate method with one of the four luciferase plasmids as well as a plasmid which expresses a DsRed fluorescent protein, which was included as a transfection control. Values for luciferase expression were then normalized to account for transfection efficiency.

The entire 5’UTR was added in front of luciferase, which includes the first exon and 11 bases of exon 2 (Figure 22A). Sequencing verified an exact match to published IRF5 5’UTRs (GenBank IDs: NM_002200.3, NM_032643.3, NM_001098627.2, and EU258897.1 for exons 1A, 1B, 1C, and 1D, respectively). Exon 1B was least-well translated, especially when compared to exon 1A (28-fold lower, p=0.048) and when compared to 1D (16-fold lower, p=0.022) (Figure 22B).
Figure 22. Different translational activity levels among the four 5’UTRs of IRF5 first exons. (A) The four 5’UTRs are shown color-coded for exon 1; exon 2 is shown in gray until the start codon, ATG, which is shown in white. (B) Four different 5’UTRs, including each first exon of IRF5 were inserted before a luciferase reporter gene, although each had the same promoter, to assay translation efficiency. These plasmids were cotransfected into HEK293T cells with a control red fluorescent protein-expressing plasmid to normalize transfection efficiency. Exon 1B was translated relatively weakly, whereas exon 1A was 28-fold higher (p=0.048) and 1D was 16-fold higher (p=0.022) by paired t-test. Bars represent standard error. Averages are from four independent experiments. RLU: relative luminescence units, RFU: relative fluorescence units, UTR: untranslated region.

2.4.5. Exon 1B translation inhibited by a hairpin

The low levels of protein made when using the 1B 5’UTR may be due to the secondary structure of the mRNA. Therefore a folding analysis was performed using mfold [151]. For IRF5 exon 1B, folding analysis predicted a hairpin which includes all of exon 1 and three bases of exon 2 (Figures 23 and 24). This hairpin was predicted when analyzing either the 5’UTR alone or the entire mRNA of IRF5 variant 2 (GenBank ID: NM_032643.3), which begins with exon 1B.

The secondary structure of mRNA is often important in translational efficiency, especially near the 5’ end, where translation initiates [191]. Therefore, DNA with the sequence of the IRF5 exon 1B was synthesized and inserted into a luciferase expression vector to investigate if the hairpin would inhibit initiation of translation.
Figure 23. 5’ hairpin in exon 1B-containing IRF5 mRNA. mRNA folding analysis was performed using mfold [151]. The numbers between bases represent the ΔG value for each base pairing. The strongest bonds are indicated with yellow boxes. The numbers outside the hairpin represent the ribonucleotide position. The 5’ end is indicated.

Within the hairpin, three G-C base pairs had the strongest ΔG value (Figure 23). At these bases, a guanosine was mutated to adenosine using custom-made oligonucleotides. These G-to-A mutations were predicted to disrupt the hairpin when full-length IRF5 variant 2 mRNA folding was analyzed (Figure 24A). The mutated exon 1B was not predicted to self anneal, but instead it annealed to distal areas of the mRNA (Figure 24B). In order to confirm that the effect of the hairpin is shape-specific and not sequence-specific, a variant was constructed with complementary changes to allow the hairpin to reform, but with an altered sequence (Figure 24C). Folding analysis also placed each 5’UTR with the luciferase gene instead of IRF5, and the hairpins were predicted to be intact in the wild type and complementary mutant versions.

The three versions of IRF5 exon 1B 5’UTRs were inserted directly upstream of the luciferase coding region in individual plasmids. Again, all were under the control of the same promoter and cotransfected into HEK293T cells with a red fluorescent protein-expressing plasmid to account
for transfection efficiency. Elimination of the hairpin structure resulted in a 5-fold increase in translation (p=0.0021). Reconstruction of the hairpin by complementary mutations abolished the increase in translation (p=0.030) (Figure 24D).

Figure 24. Inhibition of exon 1B translation efficiency by a hairpin structure. (A) Folding predictions of IRF5 variant 2 (GenBank ID: NM_032643.3) were made for wild type exon 1B using mfold software. A hairpin was predicted to form which includes all of exon 1B. (B) Three guanosine nucleotides (orange) were mutated to adenosine (green) to eliminate the hairpin structure. (C) Complementary bases were changed to uracil (blue) to re-form the hairpin, with the folding predictions shown. Numbers represent nucleotide positions; exons and start codons indicated. (D) These three versions of the IRF5 1B 5'UTR were inserted into a luciferase vector, but each had the same promoter to assay translation efficiency. These plasmids were cotransfected into HEK293T cells with a control red fluorescent protein-expressing plasmid to normalize transfection efficiency. Luciferase activity represents translational efficiency. Variant 1B_m was translated 4.8-fold higher than 1B_wt (p=0.0021) and 8.1-fold higher than 1B_cm (p=0.030) by paired t-test. Bars represent standard error; values are averages from at least six independent experiments. RLU: relative luminescence units, RFU: relative fluorescence units, UTR: untranslated region.
2.5. Five novel splice variants sequenced from IRF5 PCR amplicons

Alternative splicing is common for IRF5 mRNA, not only the first exon, but downstream as well. Splicing of IRF5 is more common in PBMCs from SLE patients, which has been attributed to increased expression of splicing components such as small nuclear ribonucleoproteins (snRNP) [165].

Reverse-transcription PCR of IRF5 was performed on LCLs generated from healthy individuals. PCR used exon 1-specific primers and a common reverse primer in exon 7. After electrophoresis, several bands were observed that were smaller than the predicted size for full-length IRF5. Bands were gel extracted, purified, and then sequenced to reveal what parts of the RNA had been spliced out. Five new splice variants were observed. They were sequenced and named variants 13 through 17 (GenBank IDs: JQ950681-JQ950685). Two variants, 13 and 14, exactly match the previously identified variant 8, except with a different first exon. All five novel splice variants exhibited skipping within exon 6. Exon 6 contains most of the nuclear export signal (NES) [144] and the instability domain [143]; also referred to as the proline, glutamic acid, serine, and threonine (PEST) domain. Variants 15, 16, and 17 change the frame and would produce a protein with an early stop codon, truncating the transactivation domain. This is similar to previously identified variants 9, 11, and 12. All newly identified variants are shown in Figure 25, with previously identified variants for comparison.
Figure 25. IRF5 mRNA variants organized by splicing events, including five novel variants. For the new splice variants shown in orange, LCLs generated from healthy volunteers were used to make cDNA libraries. PCR products were separated by size on agarose gels and sequenced. Boxes represent exons; lines are introns. Functional areas of the encoded protein are indicated at the bottom; exons at the top. The four possibilities for exon 1 are 1D, 1A, 1B, and 1C; and exon 1B is only found in cells with the rs2004640 risk T allele. The first 12 splice variants were discovered previously (blue), variants 13-17 in the current study (orange). Boxes were placed around groups based on common splicing events. The common start codon is 12bp into exon 2, except for variant 7—which lacks exon 2 and begins in frame in exon 3. Variants 9, 11, 12, and 15 to 17 include frameshift splicing events and have early stop codons (yellow asterisks). Some variants do not show sequences past the stop codon. The 3' ends of most variants are not fully sequenced, or their length is variable due to two possible polyadenylation signals. Drawn to scale, but introns were reduced in size 10:1. UTR: untranslated region, NLS: nuclear localization signal, NES: nuclear export signal.
2.6. Altered expression levels of associated genes due to the rs2004640 and CGGGG risk factors

In collaboration with the Oklahoma Medical Research Foundation, a microarray study was performed to evaluate the effect of a risk haplotype in EBV infected or non-infected B cells [192]. The haplotype included both the rs2004640 SNP and the rs10954213 SNP. The rs10954213 SNP is an A or G polymorphism. It alters the polyadenylation sequence, producing shorter, more stable mRNA in those with the risk A allele [109]. Blood samples from ten volunteers were assayed for mRNA expression in a ~25,000 gene microarray. Five IRF5 high-risk (three controls, two patients) and five IRF5 protective (two controls, three patients) gender- and ethnicity- matched volunteers were recruited to obtain samples of PBMCs for testing.

An analysis was performed to determine if any association exists between the risk haplotype and specific signaling pathways. This was done by grouping the genes by pathways which exhibited differential expression between risk and protective haplotypes. These groups were included in a global pathway analysis using the Ingenuity Pathway Analysis system. This analysis uses the curated Ingenuity Knowledge Base to associate sets of genes and expression data with established gene pathways. Fisher’s exact-test showed statistically significant association (p<0.01) in three pathways: interferon, Toll-like receptor, and B cell receptor pathways (Table 3). Each of these signaling pathways has significant implications for autoimmune disease.

For interferon pathway genes (Figure 26A), the greatest difference in expression due to the risk haplotype was an increase in the expression of interferon beta. IFNβ was 27-fold lower in risk
Table 3. Genes exhibiting differential expression between risk and protective cells in the canonical pathways identified through Ingenuity Pathway Analysis. Fold up/down column is positive in the case that the gene expression is higher in the risk haplotype cells and negative in the case that gene expression is higher in the protective cells. Akt: Ak strain transforming, CD: cluster of differentiation, Fos: feline osteosarcoma, GSK: glycogen synthase kinase, IFITM: interferon induced transmembrane, IFN: interferon, LY96: lymphocyte antigen 96, MAPK: mitogen activated protein kinase, MX1: myxovirus resistance 1, NFAT: nuclear factor of activated T cells, NFκB: nuclear factor kappa light chain enhancer of activated B cells, OAS: 2′-5′-oligoadenylate synthetase, PIK3CA: phosphoinositide-3-kinase catalytic alpha, prot.: protective haplotype, RAC: Ras-related C3 botulinum toxin substrate, STAT: signal transducer and activator of transcription, TLR: Toll-like receptor, TNFAIP: tumor necrosis factor alpha-induced protein

<table>
<thead>
<tr>
<th>Signaling Pathway</th>
<th>Gene</th>
<th>Mock infected 16 hours</th>
<th>EBV live virus infected 16 hours</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interferon</strong></td>
<td>IFNB1</td>
<td>104.95</td>
<td>3.89</td>
</tr>
<tr>
<td></td>
<td>STAT1</td>
<td>71.08</td>
<td>52.84</td>
</tr>
<tr>
<td></td>
<td>OAS1</td>
<td>56.04</td>
<td>84.55</td>
</tr>
<tr>
<td></td>
<td>MX1</td>
<td>3179.14</td>
<td>5161.02</td>
</tr>
<tr>
<td></td>
<td>IFNAR2</td>
<td>959.14</td>
<td>1637.11</td>
</tr>
<tr>
<td></td>
<td>IFITM1</td>
<td>439.88</td>
<td>944.41</td>
</tr>
<tr>
<td><strong>Toll-like Receptor</strong></td>
<td>IFNB1</td>
<td>104.95</td>
<td>3.89</td>
</tr>
<tr>
<td></td>
<td>FOS</td>
<td>116.48</td>
<td>53.84</td>
</tr>
<tr>
<td></td>
<td>MYD88</td>
<td>281.53</td>
<td>152.57</td>
</tr>
<tr>
<td></td>
<td>TNFAIP3</td>
<td>439.96</td>
<td>321.29</td>
</tr>
<tr>
<td></td>
<td>TLR1</td>
<td>26.50</td>
<td>46.76</td>
</tr>
<tr>
<td></td>
<td>LY96 (MD-2)</td>
<td>560.79</td>
<td>1274.32</td>
</tr>
<tr>
<td></td>
<td>CD14</td>
<td>401.17</td>
<td>1272.53</td>
</tr>
<tr>
<td><strong>B Cell Receptor</strong></td>
<td>CD79B</td>
<td>2222.90</td>
<td>991.48</td>
</tr>
<tr>
<td></td>
<td>CD79A</td>
<td>41.32</td>
<td>19.08</td>
</tr>
<tr>
<td></td>
<td>RAC1</td>
<td>1407.29</td>
<td>930.84</td>
</tr>
<tr>
<td></td>
<td>MAPK9</td>
<td>105.87</td>
<td>71.02</td>
</tr>
<tr>
<td></td>
<td>AKT1</td>
<td>512.79</td>
<td>353.65</td>
</tr>
<tr>
<td></td>
<td>NFKB2</td>
<td>6.44</td>
<td>4.85</td>
</tr>
<tr>
<td></td>
<td>PIK3CA</td>
<td>25.86</td>
<td>40.05</td>
</tr>
<tr>
<td></td>
<td>NFAT5</td>
<td>123.74</td>
<td>222.63</td>
</tr>
<tr>
<td></td>
<td>GSK3B</td>
<td>68.58</td>
<td>123.65</td>
</tr>
</tbody>
</table>
cells, although this difference disappeared in cells once EBV infection occurred. The mRNA for interferon-induced transmembrane protein 1 (IFITM1) was strongly expressed in the risk cells in the absence of EBV (2.1-fold difference). STAT1 was expressed at lower levels in risk cells before EBV infection (1.3-fold), but after EBV, expression increased in risk cells (1.3-fold). IFNAR2 was opposite—it was higher in the risk cells without EBV (1.7-fold), but lower in the risk cells after EBV infection (1.3-fold). IFNAR2 is one of the receptors for the IFN cytokine and thus can respond directly to IFN signaling. In both the unstimulated and EBV treated conditions, 2′-5′-oligoadenylate synthetase 1 (OAS1) was overexpressed in the risk cells (1.5-fold and 2.2-fold, respectively), as well as MX1 (1.6 and 1.8-fold).

The TLR pathway was also affected by the IRF5 risk haplotype (Figure 26B). TLR1, CD14, and LY96 were upregulated in risk cells, and TNFAIP3 was downregulated regardless of EBV presence. TNFAIP3 expression increases upon inflammation to modulate the inflammatory response, and it is crucial to limiting inflammation by turning off NFκB responses [193]. As noted above (Section 1.9.1), polymorphisms in TNFAIP3 have been associated with lupus [101-103, 194-195]. TNFAIP3 suppression has been noted in tumors, especially lymphomas [196-199]. Expression of the transcription factor Fos and the signal transducer MyD88 was lower in risk cells when infected with EBV, but higher when EBV was absent.

The B cell receptor pathway also exhibited differences in gene expression due to the IRF5 risk haplotype (Figure 26C). The B cell receptor itself is a complex of membrane-bound antibody, CD79A, and CD79B. CD79A and CD79B were downregulated 2.2 fold in the uninfected risk cells, but this difference disappeared after EBV infection. Ras-related C3 botulinum toxin
Figure 26. Altered expression of genes in the interferon, Toll-like receptor and B cell receptor pathways due to an IRF5 risk haplotype. The genes which exhibited changes in expression levels due to the IRF5 risk haplotype are shown by red to blue color mapping. The colors represent expression levels according to the included legend. (A) Interferons are produced in response to infection. Ligated interferons signal through receptors to activate JAK/STAT transcription of targets in the nucleus. (B) Toll like receptors are either extracellular or endosomal and inform cells of possible infection. They use adaptor proteins such as the MyD88 complex and alter transcription by signaling through IRFs and other transcription factors. (C) The B cell receptor (BCR) consists of a membrane bound immunoglobulin and CD79. It also consists of co-receptors such as the CD19/21/81 complex, CD45, or CD22. The Fcy receptor II (FcγRII) also feeds into BCR signaling by binding Ig molecules. Transcription is altered though MAPK, NFκB, and NFAT pathways. BLNK: B cell linker, CaM: calmodulin, CN: Calcineurin, ERK: extracellular signal regulated kinase, FCGR2B: fragment crystallizable γ receptor 2B, GRB2: Growth factor receptor-bound 2, GSK3: glycogen synthase kinase 3, IFIT3: interferon-induced with tetratricopeptide repeats 3, IFITM1: interferon induced transmembrane 1, IFN: interferon, IFNAR: interferon α receptor, IKB: inhibitor of NFκB, IKK: IκB kinase, IP3: inositol trisphosphate, IRAK: interleukin-1 receptor-associated kinase, IRF: interferon regulatory factor, JAK: Janus kinase, JNK: JUN N-terminal kinase, LBP: lipopolysaccharide binding protein, LY96 (MD-2): lymphocyte antigen 96, MAL1: mucosa associated lymphoid tissue lymphoma translocation gene 1, MYD88: myeloid differentiation primary response gene 88, MEKK: mitogen activated protein (MAP)/Erk kinase kinase, MKK: MAP kinase kinase, NEMO: NFκB essential modulator, NFAT: nuclear factor of activated T-cells, NFκB: nuclear factor κ-light-chain-enhancer of activated B cells, NIK: NFκB inducing kinase, OAS1: 2'-5'-oligoadenylate synthetase 1, PI3K: phosphatidylinositol 3-kinase, PKC: protein kinase C, PLCγ2: phospholipase Cγ2, PPARα: peroxisome proliferator-activated receptor α, RAC: Ras-related C3 botulinum toxin substrate, SHIP1: SH2-containing inositol phosphatase 1, SHP: Src homology 2 tyrosine phosphatase, SOCS: suppressor of cytokine signaling, STAT: signal transducer and activator of transcription, SYK: spleen tyrosine kinase, TAB: TAK1 binding, TAK1: transforming growth factor-β-activated kinase 1, TIRAP: Toll-interleukin 1 receptor domain adaptor protein, TLR: Toll-like receptor, TNFAIP3: tumor necrosis factor alpha-induced protein 3, TRAF6: TNF receptor-associated factor 6, TRIF: Toll-interleukin 1 receptor domain adapter-inducing interferon-β, TYK2: tyrosine kinase substrate 1 (RAC1), a gene involved in lymphocyte differentiation and survival [200] was overexpressed in risk cells under all conditions (1.5-fold in mock infected cells, 1.7 fold in EBV infected cells). The signaling protein Akt1 and the transcription factor NFκB2 were downregulated in EBV infected cells with the IRF5 risk haplotype (1.9 and 3.4 fold, respectively). In three other genes the risk haplotype led to a 1.5- to 1.8-fold increase in expression: the phosphatidylinositol 3 kinase catalytic subunit alpha (PIK3CA), nuclear factor of activated T cells 5 (NFAT5), and glycogen synthase kinase 3β (GSK3B). However, when EBV was present, the risk haplotype showed a decrease of 1.6 to 1.9 fold of the same genes (Table 3).
In a separate set of experiments, several genes were evaluated for changes in expression due to either the rs2004640 or the CGGGG risk polymorphisms. The matched sets described in Section 2.1 were treated with the TLR7 ligand imiquimod, or left untreated to see if the response differed in risk allele-containing cells compared to protective. TLR7 lies upstream of IRF5, and the experimental design was to activate IRF5 pathways with imiquimod in order to see if risk-versus-protective differences could be observed in stimulated cells that could not be observed in

Figure 27. Decreased CCR7 expression in LCLs with the IRF5 CGGGG indel 4X/4X genotype by quantitative PCR. Expression of several genes was evaluated by SYBR-green based quantitative PCR using glyceraldehyde-3 phosphate dehydrogenase as the housekeeping gene. cDNA was from imiquimod treated LCLs generated from healthy volunteers. Sample sizes are indicated in gray circles. (A) No difference in gene expression due to the rs2004640 risk allele was observed by Wilcoxon signed rank sum test (B) The CGGGG indel affected CCR7, a cell receptor involved in cytokinesis. CCR7 levels were 2-fold lower in risk cells (p=0.019) by t-test. No difference due to the CGGGG risk indel was found in other genes. BZLF1, EBNA1, LMP1, and LMP2 are Epstein-Barr virus genes; the rest are human genes. BamHI fragment Z leftward reading frame, CCR C-C chemokine receptor, EBNA: Epstein-Barr nuclear antigen, IFIT: interferon induced protein with tetratricopeptide repeats, IFN: interferon, IL: interleukin, LMP: latent membrane protein, Trim: tripartite motif
resting cells. Some EBV genes were also tested to see if the risk factors affected expression of these viral genes.

The CGGGG risk 4X polymorphism had no effect on any genes, except that CCR7 expression was lower in cells with the CGGGG risk 4X/4X genotype compared to those with the protective 3X/3X genotype (Figure 27). The rs2004640 risk T/T genotype did not affect the expression of any genes. The results in Figure 27 are from imiquimod treated cells. Genes were chosen based on their proposed differential effect when comparing IRF5 to IRF7 transcription induction. IRF7 performs a similar role to IRF5, and the two genes affect expression of some similar and some distinct genes [147]. The genes tested were the EBV genes BamHI fragment Z leftward reading frame 1 (BZLF1), EBNA1, and two latent membrane proteins (LMP), LMP1 and LMP2. mRNA expression levels were evaluated for the human genes Calreticulin, CC-chemokine receptor 7 (CCR7), interferon induced with tetratricopeptide repeats 3 (IFIT3), IFNα1, IL-6, IL-10, IRF7, Noxa, and tripartite motif 22 (Trim22).

3. Summary and discussion

The effects of the risk haplotype can be summarized in four main categories: promoter differences, apoptosis differences, mRNA differences, and downstream expression differences.

3.1. Promoter differences

The CGGGG indel is within the 1A promoter. This polymorphism has previously been shown to alter transcription factor binding. When cells have the 4X variant, an additional SP1 binding site is created. This has been shown to increase IRF5 protein levels in PBMCs [85], but decrease 1A-
specific IRF5 transcripts in thymic cells [169]. Both versions of the 1A promoter showed activity in HEK293T cells and Raji cells. As expected, SP1 sites were found in our analysis of the 1A promoter, including an extra SP1 binding site in those with the CGGGG 4X indel. SP1 is active during development, cell growth, apoptosis, differentiation, immune responses, and DNA damage responses [201].

The 1A promoter has a paired box 5 (PAX5) binding site, a gene that activates B cells at early, but not late stages of development [202]. There is an E-box, and transcription factor 12 (TCF12) is a member of the basic helix-loop-helix group of transcription factors which binds to E-boxes [203]. TCF12 was shown to bind somewhere in the promoter region of IRF5 in the ENCODE dataset [175], and the E box in 1A’s promoter is a likely site. TCF12 is known to be expressed in B cells and T cells [21]. A PU.1 site is in the 1A promoter as well; PU.1 activates gene expression during B cell development and in myeloid cells [204]. The 1A promoter showed increased activity when cells were stimulated with the TLR7 agonist imiquimod. This may be through the PU.1 site through IRF7. IRF7 is known to be activated by TLR7 [205], and PU.1 binds to a similar GAAN(N)GAA motif to IRFs. Further work is necessary to determine in which cell types or with which stimuli the 1A promoter is most active, and also if the CGGGG 4X risk variant alters this activity.

A previous report by Mancl, et al. evaluated the 1A and 1C promoters [170]. The 1A promoter was stimulated by herpes simplex, Newcastle disease, and vesicular stomatitis viruses in PBMCs, Daudi, and THP-1 cells, respectively; as evidenced by increased transcription of IRF5. A luciferase reporter gene assay in the same report also showed that IRF5’s 1A promoter is
constitutively active and contains an IRF-E consensus binding site. However, the promoter region used was different—a 596 bp region determined by a 5’ rapid amplification of cDNA ends (5’RACE) experiment. It is 939 bp upstream of the GenBank reference sequences for exon 1A and even extends past the 1D exon by 714 bp. The results of their luciferase assay cannot therefore be compared with the promoter analysis performed in this work.

The 1B promoter is not affected by the rs2004640 SNP *per se*. The promoter may function the same way in risk and protective individuals. Instead, the effect is to allow splicing of a promoted transcript onto exon 2 in those with the risk T allele. The rs2004640 SNP does therefore affect the outcome of using the 1B promoter. This is evidenced by the fact that the risk T allele at rs2004640 correlates with higher levels of IRF5 and higher levels of exon 1B usage.

*IRF5*’s 1B promoter was predicted to contain a p53 binding site. In preliminary experiments, the only promoter tested which increased in activity after inducing DNA damage was the 1B promoter. The others showed a reduction in luciferase activity, likely because of the death, cell cycle inhibition, or apoptosis of the cells due to the DNA damage. Even a mutated version of the 1B promoter, which contained an altered p53 binding site, showed a decrease in luciferase activity instead of an increase. There were also differences in apoptosis in cells with the rs2004640 risk T/T allele. These are presumed to be due to the p53 binding site in the promoter for exon 1B, since only those with the T allele can use exon 1B. This is discussed in greater detail below in Section 3.2. The 1B promoter contains SP1, IRF4, TCF12, and early B cell factor (EBF) binding sites. EBF is a B cell-specific transcription factor [206]. Further work is necessary to reveal the stimuli or cell types that use the 1B promoter.
Usage of exon 1C is lower in cells with the rs2004640 autoimmune risk factor, which is likely due to less promoter activity. The 1C promoter contains putative SP1, IRF4, and EBF sites. It was the only promoter with activator protein 1 (AP1), Myc, and STAT2 binding sites. AP1 is a heterodimer of Fos and Jun proteins, among others, which are common in immune signal transduction [207]. Myc is a proto-oncogene and is essential for B cell proliferation [208]. STAT2, when complexed with STAT1 and IRF9, is known to be activated by type I interferon [209]. The STAT2 binding sites agree with a previous report on the 1C promoter of IRF5 by Mancl, et al., which said the promoter is interferon responsive [170]. The current analysis identified the same STAT2 binding site in the 1C promoter. The difference in the two analyses is the assumed placement of the initiation site. The analysis by Mancl, et al. uses 5’RACE to determine the initiation site and they calculate the STAT2 binding site is 96bp downstream of that transcription initiation site. According to our initiation site—taken from the GenBank reference sequences which use exon 1C—the site was 47bp upstream of the initiation site, which is a more likely placement for promoter activation. Also of note, cells treated with imiquimod had lower 1C levels in proportion to the total IRF5.

Usage of exon 1D is lower in cells with the rs2004640 T allele and in cells with the CGGGG 4X allele, which is likely due to less promoter activity. The 1D promoter evaluation showed only four transcription factors’ binding sites: CTCF, IRF4, NFκB, and SP1. NFκB is a target of TLR7 [210], and thus should be activated by imiquimod treatment. This was the case, and the 1D promoter nearly tripled in usage after imiquimod treatment. The IRF5 promoter analysis also showed a CCCTC binding factor (CTCF) binding site. It is interesting that the 1D promoter has putative CTCF sites and is the furthest in the 5’ direction, since CTCF is known to block the
spread of CpG methylation by acting as an insulator [211]. This may keep the other first exons—which are downstream and have high GC content—free from heterochromatin.

3.2. Apoptosis differences
IRF5 is proapoptotic in a p53-independent manner [150, 212]. This does not preclude activation by p53, and a p53 enhancer site in exon 2 of IRF5 has been shown to activate IRF5 [213]. p53 is a main regulator of apoptosis. Exon 1B’s promoter was the only one with a putative p53 binding site, and cells with the rs2004640 risk T allele are the only cells that can use exon 1B. We therefore hypothesized that apoptosis levels would be altered in cells with the risk T allele. Homozygous risk T/T cells did exhibit altered levels of intrinsic apoptosis, but not extrinsic apoptosis, when compared to homozygous G/G protective cells. However, death receptor-induced cell death is known to be facilitated by IRF5 [126]. We did observe more living cells in rs2004640 risk T/T cells when treating with TRAIL for 24 hours, however this was not statistically significant (p=0.069). Death receptor-induced apoptosis, such as through Fas and TRAIL, signal more directly through caspases; whereas the intrinsic pathways involve mitochondrial membrane permeability before caspase induction [214]. Thus the pathways affected are likely mitochondrial permeability genes of the B cell lymphoma 2 (Bcl-2) family, although this requires further testing.

Compared to protective cells, the levels of apoptosis were higher in rs2004640 risk cells after two day etoposide treatments. This agrees with the promoter luciferase data. However the levels of apoptosis were lower in rs2004640 T/T risk cells after low concentration 48 hour etoposide treatments and after 12 hour 5FU treatments. These results appear to be contradictory. They may
be due to a threshold level of activation being reached in the large dose etoposide treatment. Also, p53 can act as both a repressor or activator of transcription depending on local factors [215]. In any case, the levels of apoptosis are affected by the rs2004640 risk T/T allele.

3.3. mRNA differences

Both the CGGGG indel and the rs2004640 polymorphisms are in the promoter region of IRF5, affect IRF5 expression, and affect first exon usage. In LCLs, the levels of IRF5 mRNA and protein were approximately doubled by these polymorphisms. The rs2004640 had the greater effect on both mRNA and protein levels. There was, however, overlap of risk genotypes—some samples had both the rs2004640 risk allele and the CGGGG risk allele—and thus repeating these experiments with a larger sample size may detect independent effects better.

The relative stability of IRF5 mRNA is not significantly affected by the rs2004640 or CGGGG polymorphisms. In addition to being quantified as a whole, IRF5 transcripts were also measured according to first exon usage. Measuring first-exon specific mRNA stability did not demonstrate that one first exon yields RNA that is more or less stable than another first exon. Instead, similar trends to those seen in Figure 20A were observed over time, with exons 1A and 1D higher and 1C low throughout. We conclude that in LCLs, IRF5 stability is not affected by the rs2004640 or CGGGG promoter polymorphisms, or by the first exon used.

In proportion to the total amount of IRF5 mRNA in a cell, the CGGGG 4X variant did little to change the balance of first exon usage, although exon 1D was used less in risk cells. Cells with the rs2004640 T allele used exons 1C and 1D less than protective cells. Taken together, this
establishes an rs2004640 splicing profile: risk allele-containing cells exclusively use exon 1B, and use less exon 1C and exon 1D. The risk rs2004640 allele increases IRF5 in spite of low translational efficiency, because it drives a two-fold increase in overall transcription. This high transcription in persons with the risk allele suggests a role for the effect of exon 1B usage on overall protein expression. An association between autoimmunity and these IRF5 polymorphisms has been established. Increased IRF5 expression is the likely source of risk.

Levels of 1A were proportional between risk and protective cells at both polymorphisms. This excludes the result that there is double total IRF5 in risk cells, therefore the total level of 1A may be higher. When measuring the total level of each exon 1 (comparing to β-glucuronidase instead of IRF5), exon 1A was used at 4.0-fold higher levels in rs2004640 risk cells, however this only approached significance (p=0.081, data not shown); the CGGGG indel exhibited a similar trend.

The spacing between these promoters may be crucial for nucleosome positioning, since transcription start sites are often nucleosome free, allowing transcription factors to bind [216]. The 1D, 1A, and 1B exons are very close on the genome (see Figure 9). The space between 1D and 1A is 276bp, and the space between 1A and 1B is 226bp. The nucleosome wraps approximately 147bp of DNA with a linker of less than 160bp [217]; thus the 1D-1A-1B spacing is sufficient for a nucleosome to bind in between each. However, they are close enough that if one transcription site is being used, the neighboring sites may be occluded. This could explain, for example, why exon 1D usage decreased when the 1A promoter can bind additional transcription factors due to the CGGGG 4X variant. Exon 1C is 2,501bp downstream from 1B and is not likely affected as directly.
IRF5 transcripts originating with exon 1B, which forms a hairpin, are translated five-fold less efficiently compared to a structure without the hairpin. This is likely due to the inability of the ribosome to bind to the occluded 5’ end of exon 1B of IRF5 and initiate translation. When compared to the translation levels of the other exon 1 transcripts of IRF5, exon 1B was translated at the lowest levels. This contradicts a previous report by Kozyrev, et al. [90], who have a much higher value for the 1B 5’UTR, although other translation levels are in agreement. This difference is likely due to their use of 5’RACE-amplified UTRs whereas our 5’UTRs used reference sequences for exons 1A through 1C and variant 12’s 5’UTR for exon 1D (see Section 2.4.4).

All of the novel 1B variants came from individuals with the rs2004640 risk T/T genotype. The two new 1C variants came from a risk T/T individual and a protective G/G individual. No effect on splicing is assumed to be due to the CGGGG polymorphism, and none were directly tested or observed. The usage of exon 1B, that is the presence of the rs2004640 T genotype, may have an effect on splicing events other than exon 1, since some splicing is linked. Examples of patterns of splicing include mutually exclusive exon use, tissue-specific exon use, and developmental-specific exon use [218-220]. However, except that only one exon 1 can be used per molecule, none of these concomitant splicing events were seen. For example, three 1B variants have exactly the same amino acid sequence as three 1A variants—1B variants 2, 10, and 13 encode the same protein isoforms as 1A variants 6, 4, and 8 respectively.

Variants 8, 9, and 12 have been shown to be constitutively nuclear [170, 221] since they lack the NES in exons 5 and 6. Although they have not been analyzed, it is likely that variants 11 and 13
through 16 are also constitutively nuclear due to similar patterns of missing amino acids. They also have an intact DNA binding domain. In fact, only one identified variant has any splicing truncation in the DNA-binding domain, variant 7. It is interesting that many of the variants can bind DNA, but have truncations in either the instability domain or the transactivation domain. If DNA binding is active but transactivation is not, the encoded protein will likely be dominant negative.

IRF5 exhibits a high degree of alternative splicing, and alternative splicing of IRF5 was shown to be higher in lupus patients [165]. The ability to alternatively splice IRF5 likely allows a fine-tuned response to a variety of immune signaling events [170], but in conjunction with autoimmune disease this high degree of alternative splicing may be detrimental. Although there are currently 17 known variants of IRF5, there are likely dozens more ways to splice IRF5. Splicing may occur in response to specific stimuli that would further explain the varied levels of first exon usage and changes in overall levels of expression.

3.4. Downstream expression differences

Several gene pathways were affected in cells with an IRF5 lupus-risk haplotype. Since IRF5 is a transcription factor, these polymorphisms likely affect B cells’ response to infection or their action in autoimmune disease. The pathway analysis placed gene expression changes in three main pathways: interferon, Toll-like receptor, and B cell receptor pathways. This pathway analysis is valuable because it allows a broader look into gene networks than simply looking at individual genes.
The interferon pathway was affected by the risk haplotype. Interferon alpha is an important cytokine in lupus etiopathogenesis [26]. The interferon alpha pathway is affected by genetic variation in the \textit{IRF5} gene, and multiple interferon signaling genes could be targets for understanding interferon dysregulation in autoimmune disease. Toll-like receptors are involved in response to infection through the recognition of pathogen-associated molecular patterns. Additionally, Toll-like receptors are important in the pathogenesis of lupus. They are capable of recognizing endogenous nucleic acids in the context of immune complexes found the sera of lupus patients, which stimulates dendritic cell maturation and IFN\(\alpha\) production [222-229], a process which also involves IRF5 itself [229]. IRF5 is an interesting transcriptional regulator in that it acts as both an activator when homodimerized with itself, and its activation is blocked when heterodimerized with IRF7 since IRF5/IRF7 heterodimers mask both transcription factors’ DNA binding domains [138, 146]. This mechanism of action may help to explain how some of these pathways can exhibit relative upregulation or downregulation depending on the other conditions in the cell. The B cell receptor pathway was also affected. The B cell receptor recognizes antigen. Stimulating a B cell through the BCR provokes the survival, maturation, and proliferation of B cells. B cells are the source of autoantibodies, and in some way the tolerance to self must be broken in order to produce these antibodies.

The pattern of gene expression observed suggests that the \textit{IRF5} risk haplotype makes these cells appear more activated in the resting state. An activated basal state would be likely to promote inappropriate cellular responses and possibly heightened sensitivity to self antigens, including those recognized by TLRs.
EBV was used in several experiments to create LCLs from PBMCs. It was used to evaluate an infected state and to see any effect on EBV gene expression due to \( IRF5 \) risk polymorphisms. EBV itself could be involved in the etiology of lupus by affecting several pathways. The three pathways identified here are all involved in EBV infection. EBV may stimulate these pathways through several mechanisms, including both infection and binding of virions to the receptors involved in these pathways.

The CGGGG indel decreased the expression of CCR7. CCR7 was first named Epstein-Barr virus induced 1, since it was found to be overexpressed upon EBV infection [230]. CCR7 is involved in immune cells migration, and EBV infection is known to alter migration [231]. Further work is necessary to determine if CCR7 levels are lower in EBV negative cells. If the risk CGGGG risk factor decreases CCR7 levels, the migration of immune cells might decrease. \( Irf5 \) knockout mice are known to exhibit large changes in migration [56-57, 232-233].

### 3.5. Conclusion

A key effect of these promoter polymorphisms, therefore, is to increase the levels of IRF5, at both the mRNA and protein levels. Several studies have shown that an rs2004640 risk haplotype causes increased expression of IRF5 mRNA or IFN-pathway genes in lupus patients [106, 153]. Our experimental design using healthy controls with and without the risk polymorphisms provided us the advantage of measuring the effect of the risk allele without the confounding factors of disease, since all of our data about promoter, apoptosis, and mRNA differences come from those with the risk factor for autoimmune disease, but not the disease itself. We thus infer a causal relationship between high IRF5 levels and autoimmune disease. This is corroborated by
The different roles of IRF5 align with the range of autoimmune diseases it is involved in. As examples, IRF5 functions as a cytokine transcription factor [143, 149] and as a p53-independent pro-apoptotic regulator [126, 212, 234]. Dysregulation of cytokine levels and apoptosis are features of SLE [122, 235-236], the disease with which the rs2004640 SNP was first associated. Also, IRF5 has been shown to be an important mediator in macrophage differentiation towards pro-inflammatory M1 macrophages [149]. With high levels of IRF5 there would be more pro-inflammatory cytokines, which would provoke a greater immune response. This has been confirmed by association of the rs2004640 risk SNP in *IRF5* with macrophage activation syndrome [237].
The etiologies of autoimmune diseases are complex, but various genetic and environmental factors contribute to their onset. Studies of genetic risk factors such as the rs2004640 SNP and CGGGG indel in IRF5 point to the pathways involved in disease and therefore to therapies which will allow for more effective treatments.

4. Material & Methods

4.1. Plasmid construction and luciferase assay

For assaying translational efficiency, oligonucleotides representing the 5’UTR with the sequence for wild type and mutated exon 1B were synthesized (Integrated DNA Technologies), annealed, and inserted into the pGL3-Promoter vector (Promega). The longer 5’UTRs which include exons 1A, 1C, and 1D were PCR amplified from cDNA using ExTaq (Takara) and the primers listed in Table 4. Each 5’UTR was ligated into the pGL3 vector. All vectors were sequenced to confirm the proper sequence. The plasmids pSIREN-DNR-DsRed or pMax-GFP (Clontech) express the DsRed and eGFP fluorescent proteins, respectively, and were used to measure transfection efficiency. HEK293T cells were cotransfected with both a luciferase plasmid and a fluorescent protein-expressing plasmid using the calcium phosphate method. LCLs were electroporated using a Nucleofector device (Lonza). The electroporation buffer was 5 mM KCl, 15 mM MgCl₂, 15 mM HEPES, 140 mM Na₂HPO₄, pH 7.2. Transfected cells were lysed and assayed for fluorescence levels before assaying luciferase activity using the Luciferase Assay System (Promega) on a Fusion αHT plate reader (Packard). Luciferase activity was evaluated in proportion to the transfection efficiency except where noted in preliminary data.
4.2. Cell lines

Peripheral blood samples were obtained from healthy volunteers after informed consent following an IRB-approved protocol. Peripheral blood mononuclear cells were isolated using lymphocyte separation medium (Mediatech). These cells were induced to form lymphoblastoid cell lines (LCLs) by incubation with Epstein-Barr virus (B95-8 strain) and 2 ng/ml cyclosporin A (Torcris Biosciences). Raji cells were from ATCC. Cell lines were maintained in RPMI (Sigma) with 10% fetal bovine serum (PerBio) with penicillin/streptomycin/amphotericin (Calbiochem) at 5% CO₂ and passaged at least weekly. HEK293T cells were a gift from Dr. Bradford Berges, and were maintained under the same conditions, but with DMEM as the base media.

4.3. Genotyping of volunteers and formation of paired samples

Genomic DNA was extracted (Qiagen) from peripheral blood mononuclear cells and genotyped using TaqMan reagents (Applied Biosystems (ABI)) on a StepOnePlus real-time PCR machine (ABI) at the rs2004640 SNP (ABI SNP Assay C9491614). Homozygous risk or protective individuals were matched by gender and ethnicity. Heterozygotes were not included in the study. Genotyping at the CGGGG indel was performed by PCR amplification and 4% agarose gel size determination. Genotyping at rs10488631 (ABI SNP Assay C2691242), which has also been shown to affect IRF5 expression [238] was also performed, Heterozygotes were not included in the study. The primers and PCR conditions are in Table 4.

4.4. Cell treatments

The Toll-like receptor 7 (TLR7) ligand imiquimod, also known as R-837, was used to activate IRF5 for some experiments. Cells were treated for 24 hours with 25 μg/ml imiquimod (InvivoGen). Actinomycin D, an inhibitor of RNA polymerase, was used to test mRNA stability. Cells were treated for 24 hours with 5 μg/ml actinomycin D (Fisher), and also analyzed at
2 hours and 6 hours [239]. cDNA preparation, quantitative PCR, primers, probes, and conditions are described in Section 4.5. UV light exposure was performed after removing media by placing the cells 75 cm from a 254 nm UV bulb for 90 seconds. Etoposide was used at 0.1 mM and 1 mM concentrations and applied for 48 hours. 5FU was used at 1.5 mg/ml, and the activating antibodies to TRAIL and Fas were used at 1 μg/ml and 5 μg/ml, respectively. All treatments used 10⁶ cells per milliliter. For apoptosis level testing, the cells were treated with 1 μg/ml PI (Sigma) and Annexin V-FITC was applied per kit instructions (Life Technologies).

4.5. cDNA libraries and PCR

cDNA preparations were made by extracting RNA using the RNaqueous system (Ambion), followed by DNase treatment (Promega), then reverse transcription using SuperScript III reverse transcriptase (Invitrogen Life Technologies). cDNA preparations were used as template for quantitative PCR. Glyceraldehyde-3 phosphate dehydrogenase (GAPDH) was used as a housekeeping control for calculating comparative CT values; however, total IRF5 was used as the control value when comparing first exon usage. SYBR green reagents (ABI) were used in all cases. In order to maintain linear relationships between CT values for first exon proportional analysis in the SYBR green-based assay, all primer sets were designed to produce amplicons of nearly equal length (size range: 98 to 109 bp, see Table 4). For IRF5 splice variant analysis Takara Taq was used. For cloning of 5’UTRs the template cDNA was used with Takara ExTaq. For cloning promoters the template genomic DNA from Section 4.3 was used with the NEB High GC PCR kit. Primers and oligos were purchased from Integrated DNA Technologies. Sequences and PCR conditions are shown below in Table 4.
### Table 4. List of primers and oligonucleotide sequences, and PCR and annealing conditions.

All are listed in 5'-to-3' orientation. Restriction enzyme cut sites or overhangs are indicated in lowercase. For quantitative PCR primers, amplicons’ lengths and %GC contents are also listed. fwd: forward primer, GAPDH: glyceraldehyde 3-phosphate dehydrogenase, IRF5: interferon regulatory factor 5, rev: reverse primer, RT: real time, UTR: untranslated region

#### Oligonucleotides:

<table>
<thead>
<tr>
<th>Oligonucleotides</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRF5 5’UTR 1B wild type luciferase top</td>
<td>agcttGTCCAGCTGCGCTGGAAAGCGAGCTCGGACCCCTCTGc</td>
</tr>
<tr>
<td>IRF5 5’UTR 1B wild type luciferase bottom</td>
<td>catggCAGAGGGGTCCGAGCTCGCTTTCCAGGCGCAGCTGGACa</td>
</tr>
<tr>
<td>IRF5 5’UTR 1B mutant luciferase top</td>
<td>agcttGTCCAACCTGGAAGCGAGCTCGAAACCCCTCTGc</td>
</tr>
<tr>
<td>IRF5 5’UTR 1B mutant luciferase bottom</td>
<td>catggCAGAGGGGTCCGAGCTCGCTTTCCAGGCGCAGCTGGACa</td>
</tr>
</tbody>
</table>

Annealed at 95°C for 1 min, 72°C for 2 min, 37°C for 2 min, 25°C for 2 min.

#### Splice variant amplification primers:

<table>
<thead>
<tr>
<th>Splice variant amplification primers</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRF5 1A fwd</td>
<td>CCTGGCGAGCCACGCAGGCAGCGCA</td>
</tr>
<tr>
<td>IRF5 1B fwd</td>
<td>GGGCCTGGAAGCGAGCTCGCG</td>
</tr>
<tr>
<td>IRF5 1C fwd</td>
<td>TAGGGAGGCTGAAACCCAAAA</td>
</tr>
<tr>
<td>IRF5 1D fwd</td>
<td>GAGGCTAGCCGGATCTGC</td>
</tr>
<tr>
<td>IRF5 common rev</td>
<td>CTTGATCTCCAGGTCGGTCA</td>
</tr>
</tbody>
</table>

Conditions: 94°C 4 min, 37 cycles of [94°C 1 min, 63.9°C 1 min, 72°C 1:30], 72°C 5 min

#### First exon 5’ UTR cloning primers:

<table>
<thead>
<tr>
<th>First exon 5’ UTR cloning primers</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRF5 1A 5’UTR fwd</td>
<td>TaagttGCCCGGCAGGTTGGCGGA</td>
</tr>
<tr>
<td>IRF5 1C 5’UTR fwd</td>
<td>TaagttGACGCCTGGAACACTCCC</td>
</tr>
<tr>
<td>IRF5 1D 5’UTR fwd</td>
<td>TaagttATTTGGCATTCCAGATTG</td>
</tr>
<tr>
<td>IRF5 common 5’UTR rev</td>
<td>CTGGAATCCAGGAGGCTCT</td>
</tr>
<tr>
<td>pGL4 sequencing fwd</td>
<td>CTAGCAAAAATGGCCTGTCCC</td>
</tr>
</tbody>
</table>

Conditions: 94°C 2 min, 40 cycles of [94°C 15 s, 60°C 1 min, 72°C 20 s], 72°C 1 min
**TaqMan-based quantitative PCR primers:**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
<th>Probe</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRF5</td>
<td>CCACCTCAGCCCTACAAGAT</td>
<td>CCACCTCAGCCCTACAAGAT</td>
<td>FAM-TCCAATGGCCTGCAC-TAMRA</td>
</tr>
<tr>
<td>IRF5</td>
<td>CTCTCTCTCTGCACCAAAAAG</td>
<td>CTCTCTCTCTGCACCAAAAAG</td>
<td>FAM-TCCAATGGCCTGCAC-TAMRA</td>
</tr>
<tr>
<td>IRF5 1A</td>
<td>ACGCAGCGACCCAGCATACCATGAGTGAAC</td>
<td>CTCCTGACCAGCATACCATGAGTGAAC</td>
<td>GAGGCTCAGCCCGGATCCGTA</td>
</tr>
<tr>
<td>IRF5 1B</td>
<td>AGGCTCGCTGGAAAGCAGC</td>
<td>AGGCTCGCTGGAAAGCAGC</td>
<td>GAGGCTCAGCCCGGATCCGTA</td>
</tr>
<tr>
<td>IRF5 1C</td>
<td>AGGCGGCACTAGCCAGGTGCAAC</td>
<td>AGGCGGCACTAGCCAGGTGCAAC</td>
<td>GAGGCTCAGCCCGGATCCGTA</td>
</tr>
<tr>
<td>IRF5 1D</td>
<td>GAGGCTCAGCCCGGATCCGTA</td>
<td>GAGGCTCAGCCCGGATCCGTA</td>
<td>GAGGCTCAGCCCGGATCCGTA</td>
</tr>
<tr>
<td>β-glucuronidase</td>
<td>FAM-CCATGAACCAGTCCATCCAGCTGCCCTCCACC-TAMRA</td>
<td>FAM-CCATGAACCAGTCCATCCAGCTGCCCTCCACC-TAMRA</td>
<td>GAGGCTCAGCCCGGATCCGTA</td>
</tr>
</tbody>
</table>

**Conditions:** 52°C, 95°C for 10min, 52 cycles of [95°C for 15s, 65°C* for 1 min] with 500 nM primers, 250 nM probe

*primer annealing temperatures were 60°C for β-glucuronidase; 65°C for IRF5, 1A, and 1B; 66°C for 1C, and 69°C for 1D

---

**SYBR green-based quantitative PCR primers:**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
<th>Probe</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRF5</td>
<td>CCACCTCAGCCCTACAAGAT (99 bp, 58.6% GC)</td>
<td>CCACCTCAGCCCTACAAGAT (99 bp, 58.6% GC)</td>
<td>CCACCTCAGCCCTACAAGAT (99 bp, 58.6% GC)</td>
</tr>
<tr>
<td>IRF5</td>
<td>CTCTCTCTCTGCACCAAAAAG</td>
<td>CTCTCTCTCTGCACCAAAAAG</td>
<td>CTCTCTCTCTGCACCAAAAAG</td>
</tr>
<tr>
<td>IRF5 1A</td>
<td>GCAGGCGGACCCAGGACGAC</td>
<td>GCAGGCGGACCCAGGACGAC</td>
<td>GCAGGCGGACCCAGGACGAC</td>
</tr>
<tr>
<td>IRF5 1C</td>
<td>CTAGGCACTGGAACCCCAAAA</td>
<td>CTAGGCACTGGAACCCCAAAA</td>
<td>CTAGGCACTGGAACCCCAAAA</td>
</tr>
<tr>
<td>IRF5 1D</td>
<td>GAGGCTCAGCCCGGATCTGC</td>
<td>GAGGCTCAGCCCGGATCTGC</td>
<td>GAGGCTCAGCCCGGATCTGC</td>
</tr>
<tr>
<td>GAPDH</td>
<td>TGCACCACCAACTGCTTAGC</td>
<td>TGCACCACCAACTGCTTAGC</td>
<td>TGCACCACCAACTGCTTAGC</td>
</tr>
<tr>
<td>GAPDH</td>
<td>GCCACCAGCCAGGGCTTCAG</td>
<td>GCCACCAGCCAGGGCTTCAG</td>
<td>GCCACCAGCCAGGGCTTCAG</td>
</tr>
</tbody>
</table>

**Conditions:** 95°C for 10 min, 40 cycles of [95°C for 15 s, 60°C for 1 min]

---

**CGGGG genotyping**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGGGG</td>
<td>GCAGGCGGAGGTACGGG</td>
<td>GCAGGCGGAGGTACGGG</td>
</tr>
<tr>
<td>CGGGG</td>
<td>GCTCTGCCCAGGGCTGGC</td>
<td>GCTCTGCCCAGGGCTGGC</td>
</tr>
</tbody>
</table>

**Conditions:** 95°C for 2 min; 12 cycles of 94°C for 30 s, 62°C for 30 s (with a decrease of 0.5°C per cycle), 72°C for 1 min; 30 cycles of 94°C for 30 s, 56°C for 30 s, and 72°C for 1 min; 72°C for 5 min.
Promoter cloning primers†:

IRF5 1A prom fwd: CTGCgctagcCAGGTCAGTGCGGGGC
IRF5 1A prom rev: CCTGagatctACTTCCCGGTCTTGCCGC
Conditions: 94°C 30 s, 40 cycles of [94°C 15 s, 62.0°C 1 min, 68°C 30 s], 68°C 5 min

IRF5 1B prom fwd: GCAGgctagcGACAGGTGGGTCCCGGC
IRF5 1B prom rev: GCAGagatctGCGGACCCCGCCCTACTCCA
Nested PCR first round: IRF5 1A prom fwd + IRF5 1B prom rev
Conditions: 94°C 30 s, 40 cycles of [94°C 15 s, 59.3°C 1 min, 68°C 30 s], 68°C 5 min
Nested PCR second round: IRF5 1B prom fwd + IRF5 1B prom rev
Conditions: 94°C 30 s, 40 cycles of [94°C 15 s, 66.0°C 1 min, 68°C 30 s], 68°C 5 min

IRF5 1C prom fwd: TAGTgctagcGCTGGTTTCCTCAGGTCTCT
IRF5 1C prom rev: CAGAagatctCAGCCCTGCCCTGGCCT
Conditions: 94°C 30 s, 40 cycles of [94°C 15 s, 60.8°C 1 min, 68°C 2 min], 68°C 5 min

IRF5 1D prom fwd: ACATgctagcCAGCTGGCTGCCTGTTGACC
IRF5 1D prom rev: TGGCagatctGTCAATTGACAACCCC
Conditions: 94°C 30 s, 40 cycles of [94°C 15 s, 59.4°C 1 min, 68°C 1 min], 68°C 5 min

pGL4 sequencing fwd: CTAGCAAAATAGGCTGTCCC
†PCR for these GC-rich promoters was performed using a High-GC kit (NEB) according to package instructions, with 10% enhancer solution included for all reactions except exon 1D.

4.6. Sequencing

For cloning of IRF5 splice variants, Takara Taq was used with LCL cDNA as template for reverse transcription PCR. After size separation on a 1.5% agarose gel, DNA was extracted from individual bands using a gel band extraction kit (Qiagen). Sequencing of IRF5 variants was performed using the forward or reverse primer used in PCR amplification. Plasmid sequencing used purified plasmid DNA and a primer upstream of the insertion site. Sequencing reactions used Big Dye terminator reagents and the 3730xl DNA analyzer (ABI). See Table 4 for primers.

4.7. Statistical analysis

An unpaired t-test or the Wilcoxon signed rank sum test was used to compare means for mRNA and protein data, as noted. A paired t-test was used for imiquimod treated versus untreated samples. Regression modeling of the polymorphisms’ effect on mRNA and protein
levels used backward selection, and included calculating the ANOVA. Paired t-test was used for luciferase levels. An alpha of 0.05 and two-tailed \( p \)-values were used in all cases. For comparing mRNA half-life slopes, regression modeling was performed for analysis of covariance.

**4.8. Computer Programs**

RNA folding analysis was performed using mfold [151] under standard settings. Alignments and viewing of sequencing files were completed using MEGA 4.0 [178]. Densitometry on western blots was calculated using Quantity One software (BioRad). An analysis of the promoters for each of the four first exons of IRF5 was performed using the ENCODE ChIP-Seq data set [175] for determining actual binding factors on the genomic region, followed by determining a consensus site using the WebLogo data in FactorBook [177]. The consensus site was then used to search the proximal promoters (~200bp upstream from the +1 sites) to encounter a proposed binding site. Consensus site screening was performed using a custom searches of ambiguous nucleotides with MEGA [178]. This involved searching using the find function, which allows for searching using the ambiguous nucleotide code. For example, a search for GAW would highlight both GAA and GAT.

**4.9. Western blotting**

IRF5 protein levels were was analyzed in untreated cells that were washed in phosphate-buffered saline and lysed in Laemmli buffer. Lysates were triturated with a 25-gauge needle and boiled. Equal amounts of whole-cell lysate were loaded onto a 10% sodium dodecyl sulfate-polyacrylamide electrophoresis gel and transferred to a nitrocellulose membrane. The membrane was probed with antibodies to IRF5 (Abcam). The secondary antibodies used were horseradish peroxidase-conjugated anti-mouse antibodies (Thermo Scientific). Protein bands were visualized
with the ECL system (General Electric). Membranes were stripped for reprobing with anti-
\( \beta \)-actin antibody (NeoMarkers) and visualized again with the ECL system.

4.10. Enzyme-linked immunosorbent assay (ELISA)

The levels of IL-10 and IL-6 were quantified using commercial ELISA kits from
eBiosciences. Plates were read on a Biotek plate reader.

4.11. Information about microarray data

The experiments which resulted in the pathway figure (Figure 26) were done in
collaboration with the Oklahoma Medical Research Foundation (OMRF); writing and data
analysis were performed by Daniel Clark at BYU. The following materials and methods section
is from the published paper [192].

Genetic testing was performed on samples obtained from the lupus family registry and
repository at the Oklahoma Medical Research Foundation. Volunteers were selected based upon
their \( IRF5 \) risk and protective haplotypes using genotypes at single nucleotide polymorphisms
rs2004640 and rs10954213. Five \( IRF5 \) high-risk (three controls, two patients) and five \( IRF5 \)
protective (two controls, three patients) gender- and ethnicity-matched individuals were
recruited. The study was approved by the institutional review board at all institutions, and
informed consent was obtained from all subjects in the study. Total cellular mRNA was purified
from lysates of infected versus mock-infected cells using the Ambion RNaqueous-Micro Kit
(ABI) according to the manufacturer’s protocol. Two-rounds of \textit{in vitro} transcription were
performed and RNA expression was analyzed on the Illumina platform using whole genome
arrays (~25,000 genes).

The microarray data were analyzed using gene set enrichment analysis and pathway
analysis to investigate changes in a gene networks. These analyses were followed by comparison
of individual gene expression differences inside these networks. Pathway analyses were generated through the use of Ingenuity Pathways Analysis (Ingenuity Systems). A data set containing gene identifiers and corresponding expression values was uploaded into the application. Each gene identifier was mapped to its corresponding gene object in the Ingenuity Pathways Knowledge Base. The expression values entered were the normalized log (intensity) values of \textit{IRF5} protective and risk haplotype individuals.

Canonical pathways analysis identified the pathways from the Ingenuity Pathways Analysis library of canonical pathways that were most significant to the data set. The significance of the association between the data set and the canonical pathway was measured in two ways: 1) A ratio of the number of genes from the data set that map to the pathway divided by the total number of genes that map to the canonical pathway; 2) Fischer’s exact-test was used to calculate a \textit{p}-value determining the probability that the association between the genes in the dataset and the canonical pathway is explained by chance alone.

\textbf{Appendix A. Effect of imiquimod}

While testing the samples from Section 2.1 above, our experimental design was to stimulate cells with imiquimod to activate IRF5 through TLR7. This imiquimod treatment was used to evaluate whether the effect of the risk polymorphisms could be seen in a stimulated state that may not have been detectable in a resting state. In addition to our experimental design of activating IRF5, the imiquimod treatments also allow us to evaluate the effect of imiquimod itself. Expression of several genes was affected by imiquimod whether measured by SYBR green-based quantitative PCR (Figure 28A and 28B) or by ELISA (Figure 28C and 28D).
Figure 28. Altered gene expression due to imiquimod treatment by PCR or ELISA. (A) Genes whose expression increased and (B) genes whose expression decreased are shown by setting the untreated sample to 1 and -1, respectively. (C) Secreted IL-10 protein levels increase upon imiquimod stimulation by ELISA. (D) Secreted IL-6 protein levels increase in a dose-dependent manner upon imiquimod treatment by ELISA. ***: $p<0.001$, **$p<0.01$ by paired t-test

One interesting result was the decrease of IRF7 levels due to imiquimod. This is unprecedented to the author’s knowledge. IRF7 is a lupus-risk gene, as noted above (Section 1.9.1). We assume that this could be due to the fact that the LCL cells used are EBV positive, and EBV’s LMP1 gene interacts with IRF7 [221]. We then performed some experiments to evaluate what would
happen to the levels of these genes in the Ramos cells and primary B cells, since they are both EBV negative. The Ramos cells decreased in IRF7 levels as well (data not shown), although they are a Burkitt’s lymphoma cell line, so they were once EBV positive and could have many changes due to the virus or the cancer state. B cells did show an increase in IRF7 levels, as is expected for TLR7 stimulation (data not shown). One possible interaction is with Epstein Barr virus induced 2 (EBI2), a human gene. EBI2 has been shown to be key to IRF7 signaling pathways [240]. Further work is needed to reveal if and how Epstein-Barr virus interferes with TLR7 and IRF7 signaling, and to determine what effect EBI2 has in these pathways.

Appendix B. Herpes simplex virus typing assay affected by region-specific phylogenetic changes

See attached publication [241]. This work was performed in collaboration with Dr. Brent Johnson, BYU, to determine why a commercial typing test failed. Upon testing, a small genetic change within herpes simplex virus 1 (HSV) glycoprotein G was determined to be the cause, changing the affinity of the typing antibody. Further analysis was performed to relate all of the then available glycoprotein G sequences phylogenetically. The phylogenetic tree was branched geographically, indicating a potential failure of this typing test in certain geographic regions.

The work is related to the other studies since it is another herpes virus, like EBV. HSV is not to the author’s knowledge, an environmental trigger for SLE like EBV is.
Characterization of herpes simplex virus clinical isolate Y3369 as a glycoprotein G variant and its bearing on virus typing

Daniel N Clark¹, Brian D Poole¹*, Daniel V Hammond¹, Tyler J Hedman¹, Danny S Catts¹, Amanda Stewart¹ and F Brent Johnson¹,²

Abstract

Background: Herpes simplex viruses exist as two major serotypes, type 1 (HSV-1) and type 2 (HSV-2). Determination of type, either HSV-1 or HSV-2, is important in accurate diagnosis and clinical control of transmission. Several tests are available for typing HSV, including a monoclonal antibody specific for glycoprotein G and several PCR assays.

Findings: A clinical isolate was identified as herpes simplex virus, but tested negative for both HSV-1 and HSV-2 antigens using type-specific monoclonal antibody assays. The isolate was determined to be HSV-1 by PCR analysis. A mutation which likely caused the monoclonal antibody non-reactivity was found in glycoprotein G. Phylogenetic analysis revealed two groups of HSV, one with the mutation and one without. Three population studies examining mutations in HSV-1 glycoprotein G were analyzed by chi-squared test. To this point, the epitope which the monoclonal antibody recognizes was only found in HSV-1 isolates from human European populations (p < 0.0001).

Conclusions: These findings suggest that the PCR-based methods for HSV typing may be more useful than the standard monoclonal antibody test in areas of the world where the variant in glycoprotein G is more prevalent.

Keywords: Herpes Simplex Virus serotyping, glycoprotein G

Findings

Herpes simplex viruses exist as two major serotypes, type 1 (HSV-1) and type 2 (HSV-2). Determination of type, either HSV-1 or HSV-2, is important in accurate diagnosis and clinical control of transmission. Tests which can determine HSV type include viral antigen tests, serological tests of human antibodies and PCR [1,2]. The importance of glycoprotein G as the test analyte is emphasized by the 2002 STD Treatment Guidelines from the CDC: “Accurate type-specific assays for HSV antibodies must be based on the HSV-specific glycoprotein G2 for the diagnosis of infection with HSV-2 and glycoprotein G1 for diagnosis of infection with HSV-1.” [3].

A clinical sample of a herpes simplex virus, designated Y3369 was isolated and proved refractory to typing. The isolate was obtained from an infected genital tract of a 48-year-old female patient. It was submitted to Richards Laboratories, Inc., Pleasant Grove, Utah, USA for diagnostic workup. The sample was incubated overnight, and then stained for virus-infected cells using a type-common polyclonal primary antibody and visualized by the immunoperoxidase technique using a rapid culture method [4,5]. The culture showed an abundance of cells positive for antibody labeling and had HSV-typical cytopathic effects, confirming the presence of HSV in the specimen (results not shown).

The Y3369 isolate was then tested using the Wampole type-specific viral antigen test for HSV glycoprotein G. A viral stock culture was generated by inoculation of a portion of the rapid culture isolate into a culture of MV1Lu cells (mink lung, ATCC CCL-64). The specimen was also incubated in C1008 cells (Vero subline, ATCC CRL-1586) and subjected to similar serotypic analysis by staining with virus-specific monoclonal antibodies (mAbs)
against HSV type 1 and type 2. These tests failed to yield a positive identification of the isolate as either HSV-1 or HSV-2 using type-specific mAb assays (Wampole Laboratories). The immunofluorescence result was negative against both reagent antisera in MV1Lu cells (Figure 1). The virus was also untypable in C1008 cells (not shown). The laboratory strains HSV-1 McIntyre and HSV-2 strain 333 were tested with mAb reagents and expected monotypic results were observed in these controls.

Determination of HSV type was done by PCR specific for the HSV pol gene using a common forward primer and type-specific reverse primers as performed by Abraham, et. al [6] and Kimura, et al. [7]. DNA was extracted (Invitrogen PureLink viral DNA/RNA mini kit) from purified virus of HSV-1 (McIntyre strain), HSV-2 (Strain 333), and from the Y3369 isolate. PCR products were then analyzed on a 1% agarose gel (Figure 2), which revealed that clinical isolate Y3369 contains the pol gene of an HSV-1 virus. To confirm the analysis, DNA was then extracted from the gel (QIAquick gel extraction kit, Qiagen) and sequenced (Parallab 350, ABI 3730xl). DNA sequencing confirmed Y3369 specimen was a strain of HSV-1 with the sequenced amplicon having 100% identity when compared to the published HSV-1 pol gene sequence (GenBank accession #X04771) and only 85% homology with the HSV-2 sequence.

Confirmation of the isolate as an HSV-1 strain was done by successful PCR amplification of HSV-1 genes UL1, UL10, UL22, glycoprotein D, and glycoprotein G (data not shown, see Table 1 for PCR conditions and primers).

Glycoprotein G was PCR amplified (see supplementary table) and sequenced. Examination of the sequences showed that the probable cause for the non-reactivity of the mAb assay was the presence of a valine residue in glycoprotein G at amino acid (AA) 111. This valine is near the immunodominant region of antibody binding during normal immune response [8]. Sequencing results were deposited [GenBank:HQ833203], and compared to other isolates on GenBank. Sequencing revealed that the clinical isolate Y3369 contains an amino acid sequence consistent with a common HSV-1 sequence found in many parts of the world [9] (Figure 3).

A meta analysis of three population studies which have sequenced this region of the HSV-1 US4 gene was conducted to determine the prevalence of valine at position 111, as was identified in our sample [9-11]. Included were isolates from individuals from China,

---

**Table 1 PCR condistions and Primers**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>UL1 for</td>
<td>5'-GAGACCCCCTCGGCTATAAA-3'</td>
</tr>
<tr>
<td>UL1 rev</td>
<td>5'-CGTTCTCGTTCTGGGTGT-3'</td>
</tr>
<tr>
<td>UL10 for</td>
<td>5'-GAGCCTTGGGCACTTATG-3'</td>
</tr>
<tr>
<td>UL10 rev</td>
<td>5'-GTGATCTGACGACAAACAA-3'</td>
</tr>
<tr>
<td>UL 22 for</td>
<td>5'-AAAACAAAGCGCTCCTCGTA-3'</td>
</tr>
<tr>
<td>UL22 rev</td>
<td>5'-GAGACACCACATGTTTTTGGG-3'</td>
</tr>
<tr>
<td>glycoprotein G for</td>
<td>5'-GCTGTGGTGGGTTGGAACACA-3'</td>
</tr>
<tr>
<td>glycoprotein G rev</td>
<td>5'-TCCCCCGCCCATACCCTAC-3'</td>
</tr>
<tr>
<td>glycoprotein D for</td>
<td>5'-TTTGTTGCGGTGTTCCGGTG-3'</td>
</tr>
<tr>
<td>glycoprotein D rev</td>
<td>5'-TCCCCATCCCCAAACCCCGACAGA-3'</td>
</tr>
</tbody>
</table>

PCR conditions for all reactions: 94°C 2 min, 40 cycles of [94°C 15 s, 57°C 15 s, 72°C 1:15], 72°C 2 min.
Japan, Kenya, South Korea, Sweden, and the United Kingdom. We discovered the valine at position 111 to be present in all HSV-1 isolates (100%, \(N = 141\)) taken from human populations from Asia and Africa. The other populations, from the UK and Sweden, contained the valine at position 111 in 36% (\(N = 185\)) of isolates (Figure 3). This valine at position 111 is located within the binding site for a commonly used typing mAb, which recognizes the epitope AFPL [10]. The phenylalanine is replaced to form the sequence AVPL in this variant.

Sequences for the middle region encoding AA 110 to 164 of glycoprotein G were analyzed and a phylogenetic tree created (Figure 4). Phylogenetic analysis groups our isolate Y3369 as an HSV-1 with sequence V (representing valine at 111) which contains the sequence AVPL instead of AFPL, as well as other common nucleotides as shown in Figure 4. All isolates from populations from Africa and Asia, as well as 36% of the European population contained the sequence AVPL, which would not be recognized by the mAb which tests for the AFPL epitope [10]. The phenylalanine is replaced to form the sequence AVPL in this variant.

Variation among HSV-1 strains that alters viral antigen assay specificity. PCR analysis is likely to succeed in HSV typing where the isolate is not recognized by the monoclonal antibody. In addition to results presented here, PCR has been used to type HSV samples on other occasions. In one study, 75 HSV-positive isolates yielded two which were untypable using type-specific antibody tests, later confirmed HSV-1 by PCR [12]. Another study yielded 1 untypable isolate of 37 tested HSV-positive isolates, which was also confirmed as HSV-1 by PCR [13]. These represent about 2% of the HSV-positive isolates in these two studies.

We have determined the presence of two phylogenetic groups of glycoprotein G. One group was only found in Europe (Figure 4, sequence F), and all the isolates in this group contain the epitope AFPL, which a common assay uses to type HSV-1. The other group (Figure 4, sequence V) was found in all tested regions, which include Africa, Asia and Europe. This group was characterized by the AVPL sequence. Y3369 is a member of this group. The two sequences differ by location statistically (\(\chi^2 = 142.8, p < 0.0001\)).
The identification of these two groups, as well as their localization to different parts of the world, may aid in developing strategies for clinical viral antigen assays for HSV typing. Although the isolates included in the meta analysis which have the AVPL sequence were not tested by us, they would likely fail to type as HSV-1 using this same test. It should be considered that tests for the viral antigen epitope AFLP be used with caution in Africa or Asia.

This variation may also alter the interaction of the virus with host. The presence of the variations in the immunodominant region of the protein suggests these mutations could be a result of viral immune evasion. These mutations may also affect the functioning of glycoprotein G, which involves attachment and entry [14]. Further tests are being performed to study what other effects this mutation has on the virus’s efficiency of infection.

**Human Subjects**

The specimen was submitted to a clinical laboratory for diagnostic workup by code number only and the work was performed under clinical laboratory licensure: Richards Laboratories of Utah, Inc., 55 East Center St., Pleasant Grove, UT 84062, Laboratory Director: Dr. F. Brent Johnson

**List of abbreviations used**

AA: amino acid; HSV: herpes simplex virus; mAb: monoclonal antibody.

**Author details**

1Department of Microbiology and Molecular Biology, Brigham Young University, Provo, UT, 84602, USA. 2Richards Laboratories, Inc., Pleasant Grove, UT 84062, USA.

**Authors’ contributions**

DC was involved in experimental design and data acquisition, performed data analysis, conducted the meta analysis, and wrote the manuscript and figures. BP designed the study, participated in data analysis, and aided in writing and reviewing the manuscript. DH designed the study, performed PCR, gels, and sequencing, and was involved in writing the manuscript. TH, DK and AS all participated in data acquisition and writing the manuscript. FBJ performed antigen testing, conceived of the study, participated in study design, and participated in manuscript writing. All authors have read and do approve the final manuscript.

**Competing interests**

The authors declare that they have no competing interests.

This work was funded by Brigham Young University.

Received: 25 May 2011 Accepted: 9 June 2011 Published: 9 June 2011

**References**


doi:10.1186/1743-422X-8-290
Cite this article as: Clark et al.: Characterization of herpes simplex virus clinical isolate Y369 as a glycoprotein G variant and its impact on typing. Virology Journal 2011 8:290.
Appendix C. Review of interferon alpha in systemic lupus erythematosus

See attached publication [26]. This work was a review article written in collaboration with members of the lab of Timothy Niewold. This work was not directly included in the dissertation, although some of the same information may be mentioned.
**Review Article**

**Interferon Alpha in Systemic Lupus Erythematosus**

*Timothy B. Niewold,1 Daniel N. Clark,2 Rafah Salloum,1 and Brian D. Poole2*

1Section of Rheumatology and Gwen Knapp Center for Lupus and Immunology Research, The University of Chicago, Chicago, IL 60637, USA
2Department of Microbiology and Molecular Biology, Brigham Young University, Provo, UT 84602, USA

Correspondence should be addressed to Brian D. Poole, bpoole@gmail.com

Received 22 January 2010; Accepted 20 May 2010

Academic Editor: Charles Via

Copyright © 2010 Timothy B. Niewold et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The pleiotropic cytokine interferon alpha is involved in multiple aspects of lupus etiology and pathogenesis. Interferon alpha is important under normal circumstances for antiviral responses and immune activation. However, heightened levels of serum interferon alpha and expression of interferon response genes are common in lupus patients. Lupus-associated autoantibodies can drive the production of interferon alpha and heightened levels of interferon interfere with immune regulation. Several genes in the pathways leading to interferon production or signaling are associated with risk for lupus. Clinical and cellular manifestations of excess interferon alpha in lupus combined with the genetic risk factors associated with interferon make this cytokine a rare bridge between genetic risk and phenotypic effects. Interferon alpha influences the clinical picture of lupus and may represent a therapeutic target. This paper provides an overview of the cellular, genetic, and clinical aspects of interferon alpha in lupus.

1. Introduction

In systemic lupus erythematosus, a finely tuned system of cells and signals is dysregulated, and the balance between tolerance and autoimmunity is disrupted. Cytokines, as a fundamental mechanism through which the immune system is kept in balance, play an important role in the etiology and pathogenesis of lupus. An example of an important cytokine involved in lupus etiology and pathogenesis is interferon alpha (IFNα).

IFNα is a pleiotropic cytokine that can affect multiple cell types involved in lupus. Several genes in the interferon pathway are associated with risk for lupus, suggesting a role for this pathway in etiology. Additionally, increased IFNα levels and expression of IFN response genes are often found in lupus. IFNα may affect the clinical manifestations of lupus and is a promising target for therapeutic interventions.

2. Cellular Aspects of IFNα in Lupus

Interferon alpha (IFNα) is a key molecule in immune regulation. It is produced by multiple cell types in response to viral infection. Plasmacytoid dendritic cells have a special role in the production of IFNα and are the main sources of serum interferon [1]. IFNα has the potential to dramatically influence the development, progression, and pathogenesis of SLE as it can influence the function and activation state of most major immune cell subsets and function as a bridge between innate and adaptive immunity.

2.1. Toll-Like Receptors and Interferon

One of the principal mechanisms through which IFNα is produced is through Toll-like receptor (TLR) signaling [2, 3]. TLR7 recognizes single-stranded RNA, culminating in interferon regulatory factor (IRF) 5 and IRF7 activation [4] and production of IFN [5–7]. Excessive TLR 7 signaling produces lupus-like autoimmunity in male Yaa mice, where an extra copy of the TLR7 gene is present on the Y chromosome [8–10]. The autoimmune phenotype conferred by the Yaa genotype is dependent on IFN α, and addition of IFNα can partially duplicate the Yaa phenotype [11]. Additionally, knocking out the IRF7 gene or inhibiting its action with pharmacologic agents inhibits antibody production against RNA-containing...
nuclear components [12], suggesting that TLR7 is essential for this type of autoantibody production.

A characteristic of many cases of lupus is the production of antibodies against RNA-containing protein complexes such as Sm, nRNP, Ro, and La. In fact, antibodies against the spliceosomal protein Sm are so specific for lupus that they are used as a diagnostic criterion. The RNA found in these complexes is capable of promoting the production of IFN-α through the stimulation of TLR7 [3, 13] (Figure 1).

Because TLR7 is located in the endosomes, RNA-containing complexes must access the interior of the cell before they are able to act as activators. Autoantibodies specific for these lupus-associated riboproteins can bind with antigens derived from apoptotic cells and form antibody-protein-RNA complexes. The Fc portions of the immune complexes are recognized and internalized by cells with Fc receptors, providing a route of entry for RNA to reach TLR7, resulting in interferon alpha production [3, 14]. This process is especially well established in PDCs [15, 16]. Interestingly, in addition to being produced as a result of TLR7 ligation, IFNα enhances TLR7 signaling in PDCs [17, 18], forming a positive feedback loop.

Despite these data and the strong association between SLE-associated autoantibodies and serum IFNα, SLE-associated autoantibodies are not sufficient for high serum IFNα in humans in vivo [19]. Healthy subjects with anti-Ro antibodies do not have high serum IFN-α, while a significant proportion of anti-Ro positive patients with SLE or Sjögren’s syndrome do have high serum IFNα, suggesting that these autoantibodies require other disease-associated factors to result in high serum IFNα in humans.

2.2. IFNα and Adaptive Immune Regulation. Excess serum IFNα and IFN-response gene expression are characteristics of lupus and are most likely the result of excessive PDC activation. Such high levels of interferon could contribute to lupus by promoting immune activation rather than tolerance. Dendritic cells are the primary activators of T cells, and affect both T-cell tolerance and activation, depending on the state of the dendritic cell. When treated with interferon alpha, dendritic cells mature and become more prone to activate T cells [20, 21]. Myeloid dendritic cells from lupus patients are able to phagocytose and present self-antigens to T cells in a stimulatory, rather than regulatory manner, a process which is interferon-dependent [22]. Such a process likely contributes to loss of T-cell tolerance to self-antigens and subsequent autoimmunity.

Exposure of the dendritic cell to IFNα contributes to T cell polarity. When CD4+ T cells are activated in the presence of IFNα-producing dendritic cells, their polarity is shifted towards IFN-γ producing cells rather than IL-4 producing cells [23, 24], which may promote autoimmunity or immune dysregulation. The T-cells activated by IFNα-treated dendritic cells also are enriched for T-follicular helper cells, a recently described cell type which are adept at activating B cells and driving antibody production [25].

Regulatory T cells (T-reg) are attracting increased attention as a mechanism of immune regulation and suppression of autoimmunity. In lupus, T-regs are often, though not always, found in lower numbers than in controls [26–31]. Those T-regs that are present in lupus are inefficient at suppressing inflammation and T-cell proliferation [27, 29, 30, 32]. T-reg development is suppressed by treatment of dendritic cells with IFNα [33]. In lupus patients, T-reg activity is diminished, due at least in part to the action of IFNα [34] indicating that increased IFNα levels in lupus patients is likely contributing to the development or maintenance of autoimmunity through suppression of T-reg cells.

B cells are important in lupus, since humoral autoimmunity is a hallmark of the disease. IFNα can prevent apoptosis and enhance proliferation of primary B cells, even in the absence of mitogenic stimuli [35]. Interestingly, isolated B cells are inhibited from developing into antibody-producing plasma cells by IFNα treatment [36]. However, this inhibition is reversed if the B cells are allowed to come into contact with monocytes, in which case IFNα actually stimulates B-cell development and antibody production [37].
The ability of IFNα to influence the activation and function of many major immune cell subsets is a testament to the wide and far-reaching effects of this cytokine. It is clear that interferon is dysregulated in lupus and that overexpression of IFNα can result from the autoantibodies present in lupus. Many components of the molecular pathways through which IFNα and TLRs drive immune activation include genetic risk factors for lupus, further implicating IFNα in lupus etiology and pathogenesis.

3. IFN and IFN-Related Genes Associated with SLE Risk

Lupus involves a combination of both environmental and genetic factors. Support for a genetic component includes a high sibling risk ratio [8–29], high heritability (greater than 66%), and higher concordance rates between monozygotic twins (20 to 40%) as compared to other full siblings and dizygotic twins (2 to 5%) [38, 39]. A large number of genetic risk factors are associated with increased susceptibility to the SLE. This genetically determined increased risk status has been referred to as a “threshold liability” [40], which is expected to be highly polygenic in nature and widely variable between individuals. Environmental factors also affect lupus susceptibility and likely interact with this “threshold liability”, but as in the case of genetic factors, there is no single environmental cause. A person may have only a few of the genetic risk variations and never get SLE despite exposure to environmental triggers. In contrast, another person may have many of these variations and then develop SLE on first exposure to an environmental trigger.

3.1. Lupus-Associated Risk Loci. Research into the etiopathogenesis of SLE has recently been advanced by several large scale case-control genetic studies, including genome-wide association scans. There is now a pool of approximately 30 genes that have been associated with SLE susceptibility with a high degree of statistical certainty and many others with probable evidence for association (reviewed in [41–45]). With this large number of SLE-associated genes, we can begin to group the list of identified SLE associated genes which should provide insight into initial disease pathogenesis into functional categories. These categories include TLR and IFN signaling, apoptosis and clearance of immune complexes, and B- and T-cell signaling. Several genes affecting the interferon pathway have been associated with risk for lupus. The Interferon pathway normally serves an important function in defense against viral infection. Yet in people with genetic predisposition, environmental triggers such as viral infections may tip the scales in favor of autoimmunity.

Once a genetic variation is identified, functional inference then characterization is necessary to move from identification to an understanding of how the variation affects the etiology or pathogenesis of SLE. Since most of the genes involved in genetic susceptibility to SLE have been identified only recently, there remains much work to identify the functional differences in the genetic associations. However, work done thus far in human cohorts is promising, and the categories of genes and loci associated with risk of lupus already suggest pathways that are of high importance.

3.2. Interferon Regulatory Factors. Certain lupus-associated genetic variations have been shown to directly increase IFNα levels or response to IFNα signaling. Interferon regulatory factor 5 (IRF5) has been confirmed as a risk locus in several different ethnic groups [46–50]. Three main functional variants in IRF5 have been described, which combine to form a risk haplotype in individuals of European descent [51]. One of these loci, at rs2004640, creates an alternate splice site (exon 1B) in the untranslated first exon. Another is a copy number variation of a 30-bp insertion/deletion sequence in exon 6, and the final is rs10954213, which creates an alternate polyadenylation site, resulting in shorter, stabler mRNA [52].

Since IRF5 activates IFNα production, these more stable variants may pose a risk due to their ability to produce excess IFNα. In fact, studies of this gene in human SLE cohorts have shown that the risk variant predisposes to greater serum IFN-α, supporting the idea that the risk haplotype is a gain-of-function variant [53]. IRF5 itself is activated by IFNα signaling, producing a potential positive feedback loop. Another IRF, IRF7, has been highlighted by the association of the IRF7/KIAA1542 locus with lupus in recent studies [54, 55]. Several SNPs in this area were shown to correlate with IFNα levels and alter autoantibody profiles in certain ethnicities [56].

IRF5 and IRF7 are activated by signaling through the endosomal toll-like receptors (TLRs) 7, 8, and 9. Interestingly, both of the IRF variants which are implicated in SLE predispose to higher serum IFN-α, but only in the presence of SLE-associated autoantibodies [53, 56] suggesting that these autoantibodies may provide chronic stimulation of the endosomal TLR pathway of IFN-α generation that when combined with gain-of-function polymorphisms in the IRFs results in dysregulation of the pathway in vivo. Additionally, TLRs 8 and 9 were identified in recent studies as containing susceptibility loci to SLE [57, 58]. The role of TLRs in the interferon production was discussed above.

3.3. Interferon-Associated Genes. Another confirmed locus of susceptibility is in the gene encoding IL1 receptor-associated kinase 1 (IRAK1). This kinase is part of the signal transduction which follows TLR ligation. In a mouse model of lupus, IRAK deficiency eliminated most lupus symptoms, showing the importance of this key intermediate [59]. Since this gene is on the X chromosome, gene dosage may contribute to the risk and the prevalence of the disease in women [59].

Two interacting proteins involved in inflammation, TNFa-induced protein 3 (TNFAIP3) and TNFAIP3-interacting protein 1 (TNIP1), have been identified as risk loci [60–64]. TNFAIP3 encodes the protein A20, which helps turn off signaling through NFκB after an inflammatory response [65, 66]. TNIP1 interacts with TNFAIP3 and is involved in several signal transduction pathways.
4. Clinical Aspects of IFNα in Lupus

Lupus primarily affects women in the reproductive years; however people of all ages, genders, and ancestral backgrounds are susceptible. Disease features range from mild manifestations such as rash or arthritis to life-threatening end-organ manifestations such as glomerulonephritis or thrombosis, and it is difficult to predict which manifestations will affect a given patient.

4.1. IFN-α as a Causal Factor in Human Lupus. A number of patients treated with IFNα have developed lupus or lupus-like syndrome [72–74]. In these reports, many specific manifestations of idiopathic lupus such as malar or discoid rash, oral ulcers, photosensitivity, renal involvement, and anti-Sm and anti-dsDNA antibodies were represented, suggesting that these cases were not “drug-induced” SLE but instead resembled idiopathic SLE [73]. Discontinuation of IFNα typically resulted in remission of SLE symptoms [73], supporting a causal relationship with IFN-α. While only a minority of patients treated with IFNα develop SLE (<1% of patients) [75], these data support the idea that IFNα can be sufficient to induce SLE in some individuals. Many more IFNα-treated individuals develop a “lupus-like” syndrome [74], with some SLE symptoms which are insufficient to meet formal diagnostic criteria for SLE [76]. IFN-induced SLE can be severe, and there are reports of life-threatening multiorgan SLE involvement including glomerulonephritis, serositis, discoid rash, myopericarditis, and vasculitis [77, 78].

Another finding which supports the hypothesis of IFNα as a primary causal factor in human SLE is the clustering of high serum IFNα in lupus families [79]. Patients with lupus and their healthy relatives have higher serum IFNα activity as compared to healthy unrelated individuals [79]. Strong familial correlations in serum IFNα were observed regardless of disease status (affecteds versus unaffecteds), and SLE probands in the same family tended to have similar IFNα levels [79]. Spouses of SLE patients did not have high serum IFNα activity, and taken together these data suggest that high serum IFNα is a heritable risk factor for SLE. Interestingly, age-related patterns of serum IFNα were also observed in SLE families in which the ages of highest IFNα mirrored the ages of peak SLE incidence [80, 81]. The discovery of several lupus risk loci in IFN-related genes provides further support for the above observation that serum IFN-α is heritable, and the SLE risk variants of each of these genes result in a gain of function increase in IFNα signaling as detailed above.

4.2. Clinical Correlations with IFN Alpha. A very strong correlation is consistently observed between the presence of SLE-associated autoantibodies, such as anti-Ro, anti-La, anti-Sm, anti-RNP, and anti-dsDNA [79, 82]. Lupus patients with high serum IFNα had a significantly higher prevalence of cutaneous and renal disease in most studies [82–84]. It is interesting that both of these clinical manifestations share an association with a particular serology (rash with anti-Ro and nephritis with anti-dsDNA), and whether these clinical manifestations are associated independent of serology has not been shown to our knowledge.

A number of studies have shown that IFNα correlates with disease activity when assessed cross sectionally [82–85]. Results are conflicting regarding the potential fluctuation of IFNα with disease activity in SLE, and there are a number of studies which did not find a longitudinal correlation [86, 87]. In these studies, a cross-sectional relationship between IFNα and disease activity is still observed, suggesting that IFNα may indicate those patients who generally have higher
disease activity as compared to other patients. A recent prospective study evaluated the utility of serum interferon-regulated chemokine levels as potential biomarkers of SLE disease activity [88]. In this study, IFNα-induced chemokines correlated with disease activity cross-sectionally, rose at the time of a flare, and decreased as the disease remitted [88]. In this same study, high chemokine levels were predictive of SLE flare over the next year in a subset of patients.

4.3. Anti-IFNα Therapies in Lupus. Given all of the studies presented above, there has been considerable interest in therapies which block IFNα. To date there is one published study describing a phase I trial of a fully human monoclonal antibody that binds to the majority of the subtypes of human interferon alpha [89]. Treatment with this anti-IFN antibody resulted in a dose-dependent inhibition of interferon-induced gene expression in peripheral blood cells as well as skin lesions in patients with mild to moderate SLE [89]. No obvious safety signals were reported during the phase I trial of anti-IFN therapy, and the proof-of-principle analyses supported a biological effect blocking the IFN pathway in humans. Phase two trials to assess efficacy of these agents in treating SLE are currently underway. There are many known predictors of high serum IFNα in SLE patients, including both serologic and genetic markers outlined in this paper. We anticipate that incorporation of these variables into clinical trial design would enhance efficacy and potentially minimize side effects by targeting the most relevant patient group. Long-term safety data will be important, since IFNα is such a highly conserved and important immunological mediator of viral defense.

5. Conclusions

IFN-α is associated with SLE through multiple lines of evidence. These include genetic, immunological/serological, and clinical associations, as described in this review. It is likely that IFN-α plays a key role in SLE etiology, pathogenesis, and/or disease persistence. Despite this large body of evidence associating IFN-alpha with lupus, the association between interferon alpha and SLE is largely inferential. The exact cellular and immunological mechanisms through which IFN affects lupus also remain undiscovered for the most part. These mechanisms and pathways are potentially fertile areas for future investigation. Such studies will likely lead to new therapeutic targets as well as a greater understanding of lupus as a disease.

References


Appendix D. Review of cytokine inhibitors for the treatment of systemic lupus erythematosus

See attached publication [242]. This review of current literature was written with members of the Poole laboratory. This copyrighted work was summarized, and an alternate figure was included in this dissertation (Section 1.8.1).
REVIEW

Cytokine inhibition as a strategy for treating systemic lupus erythematosus

Daniel N. Clark, Jillian L. Markham, Chad S. Sloan, Brian D. Poole*  

Department of Microbiology and Molecular Biology, Brigham Young University, 857 WIDB, Provo, UT 84602, USA  

Received 2 September 2012; accepted with revision 4 November 2012

Abstract Cytokines regulate and control the immune system. In systemic lupus erythematosus, several of these cytokines are overexpressed and contribute to the pathogenesis of the disease. Cytokine inhibition has been successfully used to treat other rheumatic and autoimmune diseases, and several cytokines are currently being investigated to determine whether inhibition would be therapeutic in lupus. The cytokines discussed in this review have all undergone clinical trials, and include TNF-α, IL-1, IL-6, IL-10, IL-15, IL-17, IL-18 and IL-23. Inhibition of the majority of these targets was safe and showed some efficacy in treating lupus. Cytokine inhibition strategies have just started to realize their potential for the treatment of this difficult disease, and show great promise for the future.

© 2012 Elsevier Inc. All rights reserved.

Contents

1. Introduction .......................................................... 0
2. Current cytokine inhibitors undergoing trials as lupus treatments .................................................. 0
   2.1. Interleukin 6 ...................................................... 0
   2.2. Tumor necrosis factor-α ................................................ 0
   2.3. Interleukin 10 ..................................................... 0
   2.4. Interleukin 1 ...................................................... 0
3. Potential future cytokine targets .............................................. 0
   3.1. Interleukin 17A ..................................................... 0
   3.2. Interleukin 23 ..................................................... 0
   3.3. Interleukin 15 ..................................................... 0
   3.4. Interleukin 18 ..................................................... 0
4. Conclusions .......................................................... 0

* Corresponding author at: 857 WIDB, Provo, UT 84602, USA.
E-mail address: Brian_poole@byu.edu (B.D. Poole).

1521-6616/S - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.clim.2012.11.001

Please cite this article as: D.N. Clark, et al., Cytokine inhibition as a strategy for treating systemic lupus erythematosus, Clin. Immunol. (2012), http://dx.doi.org/10.1016/j.clim.2012.11.001
1. Introduction

Systemic lupus erythematosus (SLE) is a highly complex and diverse disease. Despite substantial research, the etiology and pathogenesis of lupus are still not well understood. Additionally, the course of lupus, with its remission and flares, makes evaluation of therapies difficult. Due to these obstacles, treatment for lupus has lagged behind many other diseases, with only one new medication for lupus treatment approved by the United States Food and Drug Administration (FDA) in the last 50 years.

Infusion with monoclonal antibodies or other specific proteins, such as soluble receptors that specifically target and inhibit cytokines is revolutionizing the practice of rheumatology. These agents have been used with good effect in the treatment of rheumatoid arthritis and other autoimmune or inflammatory disorders such as Crohn’s disease and psoriasis. Given the importance of inflammation and immune control in lupus, these agents hold great promise for treating lupus as well. Inhibition of several cytokines as therapeutics for lupus is under investigation, and one has met with success and FDA approval. However, given the complexity and course of lupus, the results of these trials have been somewhat mixed. This review will summarize the current approaches and strategies for inhibiting cytokines as a therapeutic mechanism in the treatment of lupus, and will include a discussion of cytokine inhibitors that have been tested in lupus as well as those which are applicable to lupus and have been tested in other diseases.

2. Current cytokine inhibitors undergoing trials as lupus treatments

Two of the most promising cytokines that have been inhibited as a treatment for lupus, BlyS and IFN-α, will be discussed in other articles in this special issue, therefore they will not be included here. Instead, the cytokines IL-6, TNF-α, IL-1 and IL-10, inhibitors of which have undergone clinical trials for lupus, will be discussed.

2.1. Interleukin 6

IL-6 is a cytokine produced by many cell types. It has multiple effects on many target cells, inducing CD4+ T cell differentiation, B cell development, and the production of acute phase proteins. It also drives production of IL-17-producing T cells [1]. IL-6 is found at increased levels in lupus patients compared to controls [1], and is also higher in lupus patients with nephritis compared to either controls or lupus patients without renal involvement [2]. B cell-produced IL-6 has been shown to contribute to autoantibody production [3]. Certain IL-6 promoter polymorphisms may contribute to genetic risk for lupus [4].

IL-6 deficient mice are resistant to lupus. IL-6-deficient MRL-Fas(lpr) mice have delayed onset nephritis and much higher survival than control mice, along with decreased cellular infiltration, complement deposition, and Ig deposition [5]. Other IL-6 deficient mice showed that anti-DNA antibody production was dependent on IL-6 in pristane-induced lupus, although the development of antibodies against RNA-binding proteins was not [6], suggesting different pathways for autoantibody production in lupus (Fig. 1).

An anti-IL-6 receptor antibody, Tocilizumab, is approved for use in rheumatoid arthritis, with seven phase three trials completed [7]. The safety profile and effectiveness of IL-6 blockade in rheumatoid arthritis is therefore well-established. There has been one successful open-label phase I dose-escalation trial of Tocilizumab in SLE [8]. The major side effect of treatment was neutropenia, with 56% of participants experiencing neutropenia at the highest dose (8 mg/kg). One participant was withdrawn because of neutropenia, however, no neutropenia-related infections were identified [8]. Neutropenia was also noted in the studies of Tocilizumab for rheumatoid arthritis, but in those studies was also not associated with infection or malignancy [7], although there is a higher risk of infection with Tocilizumab treatment (Table 1).

Tocilizumab showed promise in treating lupus, with effects that seemed directed at autoantibody production. The modified SELENA-SLEDAI scores decreased moderately but significantly from a mean of 9.5 to 5.5 [8], with most of the improvement in rash and arthritis. Anti-dsDNA levels decreased by a mean of 46.8%. This decrease in autoantibodies may be associated with decreased circulating plasma cells. Circulating plasma cells decreased by nearly 36% in the treated individuals, and remained at this low level during follow up [8]. It may be the case that the drop in plasma cells and therefore autoantibodies is responsible for the decreased rash and arthritis in treated volunteers, but these clinical responses may also be due to some other aspect of IL-6 blockade.

2.2. Tumor necrosis factor-α

TNF-α is a proinflammatory cytokine with pleiotropic effects on multiple cell types. TNF-α activates macrophages, induces the release of further proinflammatory cytokines, regulates apoptosis of lymphocytes and other cells, and aids in cell migration [9]. In lupus, TNF-α acts as a growth factor for B cells stimulating production of IL-6 and IL-1. NZB/W mice with low expression of TNF-α develop severe lupus-like disease, but addition of TNF-α later in disease also exacerbates lupus [10-12]. These results suggest that TNF-α aids in preventing the development of lupus, but once established, worsens the resulting inflammation and pathogenesis.

Inhibition of TNF-α has met with substantial success in treating rheumatoid arthritis, as well as other inflammation-mediated diseases such as Crohn’s disease and spondyloarthritis. Lupus would seem at first glance to be a good candidate for TNF-α inhibition, since TNF-α is significantly increased in the...
serum of lupus patients [2,13,14]. Some initial studies showed that treatment of lupus patients with TNF-α inhibitors had a therapeutic effect, especially in combination with other agents. However, inhibition of TNF-α can also have a stimulatory effect on lupus. Treatment with TNF-inhibitors can induce lupus-like autoimmunity, which in some cases may be severe enough to be diagnosed as lupus. After discontinuation of the anti-TNF-α treatment, the lupus-like autoimmunity normally resolves [15–17]. Also, switching to a different TNF-α inhibitor often resolves the lupus-like autoimmunity [18,19].

The reason why TNF-α inhibitors induce lupus-like autoimmunity is not clear. One explanation is that TNF-α is necessary at a threshold level to suppress the production of Interferon-α (IFN-α) [20–22]. Therefore, when TNF-α levels decrease, more IFN-α may be available to drive lupus-like autoimmunity.

Dysregulation of apoptosis and clearance of apoptotic cells also likely contributes to the development of lupus-like disease after TNF-α inhibition. Anti-TNF-α antibodies inhibit proteins, such as amyloid P and complement components C1q and C4b, that aid in the clearance of apoptotic cells [23–27]. Another possibility for the mechanism of TNF-inhibitor-induced lupus is the altered apoptosis of mature lymphocyte populations [28], contributing to excessive apoptotic debris [29]. Genetic predisposition is also involved, with lupus-prone mice exhibiting altered TNF-α receptor signaling molecules. This leads to loss of apoptotic control of autoreactive cells. Decreased apoptosis of these cells is likely to contribute to escape from tolerance of autoreactive lymphocytes in the absence of the apoptotic signal produced by TNF-α [29,30].

Despite the ability of TNF-α inhibitors to induce lupus-like syndrome, there are currently multiple clinical trials underway investigating the use of TNF-α inhibitors for the treatment of lupus. For example, two studies investigating the use of Infliximab to treat lupus patients showed significant improvement in SLEDAI scores, and had no significant reported safety concerns, although these studies only examined 27 and 6 lupus patients, respectively [31,32]. Increased anti-nuclear antibodies were observed in one of the treated groups in one study [32]. A phase II trial of Etanercept in treating lupus was terminated. The variable and somewhat contradictory results of these studies mean that TNF-α inhibitors remain a complicated but potentially important part of the future of lupus treatment.

### 2.3. Interleukin 10

IL-10 is a somewhat contradictory cytokine. It is generally considered an anti-inflammatory and immunosuppressive cytokine. However, it is overexpressed in lupus and in some cases acts as a lupus-promoting molecule. When considering IL-10 as a potential target for inhibition, therefore, it is important to balance its role in promoting lupus with its general role in regulating the immune system.

Multiple cell types produce IL-10. These include Type 1 regulatory T (TR1) cells, some FoxP3+ regulatory T cells, macrophages, myeloid dendritic cells, and B cells, as well as many other cell types [33]. IL-10 produced by TR1 cells functions primarily to reduce inflammation and control immunopathogenesis [33,34]. IL-10 suppresses the antigen-presenting and T-cell stimulatory capacity of macrophages and dendritic cells, limiting and controlling the subsequent T cell response [35].

Elimination of IL-10 in mouse models can lead to excessive inflammation. IL-10 knockout mice develop inflammatory colitis [36]. Knocking out IL-10 in the lupus model MRL-Fas(lpr) mice led to faster disease development, more severe disease, and more mortality than in IL-10 competent mice [37].

In contrast to the suppressive activities of IL-10, on antigen presenting cells especially, IL-10 can have substantial activating

<table>
<thead>
<tr>
<th>Cytokine Inhibited</th>
<th>Drug Name</th>
<th>Current Phase of Clinical Trial</th>
<th>Therapeutic Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>TNF-α</td>
<td>Infliximab</td>
<td>Phase I complete (lupus)</td>
<td>SLEDAI score improvement</td>
</tr>
<tr>
<td></td>
<td>Adalimumab</td>
<td></td>
<td>Worsened autoantibody levels</td>
</tr>
<tr>
<td></td>
<td>Certolizumab</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Golimumab</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Etanercept</td>
<td>Phase II terminated (lupus)</td>
<td></td>
</tr>
<tr>
<td>IL-1</td>
<td>Anakinra</td>
<td>FDA approved (Rheumatoid arthritis)</td>
<td>Decreased muscle pain</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Decreased Arthritis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Variable complement levels</td>
</tr>
<tr>
<td>Canakinumab</td>
<td>FDA approved (CAPS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rilonacept</td>
<td>FDA approved (CAPS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL-6</td>
<td>Tocilizumab</td>
<td>Phase II</td>
<td>SLEDAI score improvement</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Decreased autoantibody levels</td>
</tr>
<tr>
<td>IL-10</td>
<td>B-N10</td>
<td>Phase I equivalent complete</td>
<td>Improved M-SLEDAI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Decreased steroid use</td>
</tr>
<tr>
<td>IL-15</td>
<td>HuMax-IL15/AMG 714</td>
<td>Phase I/II complete (Rheumatoid arthritis)</td>
<td>N/A for lupus</td>
</tr>
<tr>
<td>IL-17</td>
<td>Ixekizumab</td>
<td>Phase III complete (Psoriasis)</td>
<td>N/A for lupus</td>
</tr>
<tr>
<td></td>
<td>Secukinumab</td>
<td>Phase III complete (Psoriasis)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Brodalumab</td>
<td>Phase II complete (Psoriasis)</td>
<td></td>
</tr>
<tr>
<td>IL-18</td>
<td>R h IL-18BP</td>
<td>Phase I complete (Rheumatoid arthritis and Psoriasis)</td>
<td>N/A for lupus</td>
</tr>
<tr>
<td></td>
<td>GSK1070806</td>
<td>Phase III complete (Psoriasis)</td>
<td>N/A for lupus</td>
</tr>
<tr>
<td>IL-23</td>
<td>Ustekinumab</td>
<td>Phase III complete (Psoriasis)</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1** Cytokine inhibition-based therapies and their effects on lupus.
effects on other cell types. Human B cells are especially prone to IL-10 stimulation, where it acts as a potent growth factor and stimulates production of immunoglobulin [38]. IL-10 also induces class switching and differentiation into plasma cells [39]. The finding that IL-10 is present in higher concentrations in lupus patients than controls, and is still higher in lupus patients with active disease compared to inactive, supports the idea that IL-10 contributes to lupus [40,41].

Inhibition of IL-10 has been tried in vitro, in mouse models, and in human lupus patients. Inhibition of IL-10 in human serum from lupus patients prevents apoptosis from being induced by that serum, indicating that the excess IL-10 in lupus patient serum may contribute to apoptosis [42]. SCID mice injected with PBMCs from human lupus patients produced less IgG when treated with anti-IL-10 antibodies [43], and similar mice showed less renal impairment after IL-10 blockade [44].

One clinical trial of IL-10 blocking antibodies as a therapy for lupus was reported [45]. The antibody was administered to six lupus patients over a period of 21 days, and followed up at six months. One patient developed chills attributable to the treatment, which was stopped for that patient. No other treatment-related adverse events were noted. Disease activity was measured by the Mexican SLEDAI test. Scores on this test improved substantially for all patients during treatment. All patients also decreased their dose of steroids during treatment due to improvement in symptoms [45]. However, the lack of a control group, as well as the limited number of patients requires that these results be interpreted cautiously.

IL-10 remains an important and somewhat contradictory cytokine. While IL-10 blockade is being considered as a lupus therapeutic, at the same time recombinant IL-10 to increase IL-10 levels is also being studied. Well-designed experiments are necessary to uncover the potential of IL-10 modulation and conditions wherein it would be useful as a treatment for lupus.

2.4. Interleukin 1

The main role of IL-1 is to promote inflammation [46]. There are 11 members of the IL-1 family, which includes both cytokines and inhibitors. Cytokines in this family that are targets for inhibition in lupus include IL-1α, IL-1β, and IL-18 [46]. IL-1β and IL-1α bind the same receptor. However, IL-1β is normally secreted, while IL-1α is predominantly cell-bound. IL-1 exerts its effects on inflammation by several mechanisms. It is a costimulator for T-cells [47], and acts to enhance the generation of Th17 cells [48]. IL-1 also acts to enhance migration of inflammatory cells and to stimulate production of other proinflammatory molecules, such as prostaglandin E2 [47]. IL-1 receptor antagonist (IL-1RA) is a naturally occurring protein that antagonistically competes with both IL-1α and IL-1β, thus blocking their function. The ratio between IL-1 and IL-1RA is important in balancing the level of inflammation and immune activation.

IL-1 has been associated with lupus through correlation between IL-1 and IL-1RA levels and lupus activity, as well as IL-1 genetic polymorphisms. IL-1 concentration is higher in lupus patients than controls, and higher in lupus nephritis than lupus without nephritis [49]. Levels of IL-1RA were found to correlate with flare and kidney involvement in lupus patients [50,51]. There are specific polymorphisms in the IL-1 gene cluster that are associated with risk for lupus [52,53]. Additionally, mice with IL-1β knocked out are resistant to the development of lupus induced by injection of anti-DNA antibodies, though mice lacking IL-1α did not receive this protection [54].
3. Potential future cytokine targets

IL-17, IL-23, IL-15 and IL-18 are potential targets for anti-cytokine therapies for SLE. Some have already shown promising results in clinical trials for rheumatoid arthritis, psoriasis, and other diseases. These interleukins are involved in important signaling pathways such as B cell activation, inflammation, apoptosis and the T cell response.

3.1. Interleukin 17A

The cytokine IL-17A is overexpressed in lupus [58]. IL-17A exerts a multitude of effects, including inducing IL-6 production, cell recruitment to inflamed locations, and differentiation of inflammatory cells and B cells [59] (Reviewed in [60,61]). One of the most important effects of IL-17A in terms of lupus is inhibiting the differentiation of regulatory T cells. [62]. IL-17A is primarily produced by Th17 cells and CD4+, CD8-double negative T cells. These cells differentiate in the presence of TGF-β and IL-6. In the absence of IL-6, the naïve T cells tend to instead differentiate into regulatory T cells, which suppress autoimmune responses [63]. IL-23, produced by antigen presenting cells, is necessary for the growth and maintenance of IL-17 producing cells [64].

IL-17A is an important contributor to lupus. Lupus patients have increased numbers of IL-17-producing cells [65], as well as IL-17A concentrations, which correlate with disease activity [66]. Multiple mouse models of lupus show high IL-17A concentrations [67]. The MRL-Faslpr lupus mouse model shows high levels of IL-17A, as well as increasing levels of the IL-17A and IL-23 receptors as the disease progresses [68].

Because of the complicated layers of cytokines responsible for driving IL-17A production and the maturation and maintenance of IL-17-producing cells, there are multiple potential therapeutic targets that would inhibit IL-17A production. Some of these strategies include directly inhibiting IL-17A, inhibiting IL-23, inhibiting IL-6, and blocking JAK/STAT signaling, which is necessary for IL-17A production [60].

An IL-17A inhibitory monoclonal antibody, Ixekizumab, has undergone clinical trials for psoriasis and rheumatoid arthritis. The safety profile appeared good, with no severe adverse effects during a phase II trial for psoriasis or a phase I trial for rheumatoid arthritis [69,70]. Secukinumab, another anti-IL-17A monoclonal antibody, has been tested for safety and efficacy in several autoimmune diseases. One study, examining Crohn’s disease, found that the treatment group in this study had significantly more infections, especially fungal infections, than the control group [71]. Other studies examining the use of Secukinumab in psoriasis and other inflammatory diseases have had good safety profiles [72,73]. An additional antibody, Brodalumab, targets the IL-17 receptor, rather than the cytokine itself, and has also had good response and safety results in studies of psoriasis [74].

3.2. Interleukin 23

Inhibiting IL-23 would likely have many of the same effects as inhibiting IL-17A, since IL-23 is necessary for the expansion and maintenance of IL-17-producing cells. IL-23 receptor-deficient mice are resistant to lupus, with fewer IL-17 producing cells, autoantibodies, and symptoms [75]. There is currently an anti-IL-23/IL-12 antibody, Ustekinumab, in use for treating psoriasis and Crohn’s disease that does not respond to Infliximab [76]. No trials have been performed using Ustekinumab for treatment of lupus, although there are case reports of Ustekinumab being used for psoriasis also being beneficial for cutaneous lupus [77,78].

3.3. Interleukin 15

IL-15 is a pleiotropic cytokine that is produced by multiple cells and has different effects depending on the target cell. It is important for the activation and survival of natural killer cells and CD8+ memory T cells [79]. Phagocytic, antigen-presenting, and proinflammatory activities of macrophages and dendritic cells are activated by IL-15 [80]. Importantly for the development of autoimmune disease, IL-15 enhances the activation and maintenance of IL-17-producing T cells, which have been discussed above [81].

IL-15 has an interesting signaling mechanism, in that it is capable of being recycled by cells with the IL-15Rα chain and presented to neighboring cells. In this way, IL-15 functions primarily in membrane-bound form through cell-cell contact [80,82].

IL-15 has been associated with lupus through both patient and animal studies. Lupus patients have high serum concentrations of IL-15 compared to controls [83–86], although this finding is not universal [87]. Lymphocytes from lupus patients do not respond as strongly to IL-15 as do controls [83]. The membrane-associated form of IL-15 was observed to be overexpressed in the male BXSB mouse model of lupus [84].

Collagen-induced arthritis can be ameliorated or prevented when IL-15 signaling is inhibited through the use of an IL-15-Fc fusion protein (CRB-15) that targets IL-15 receptor-expressing cells in mice [88]. Blockade with a different IL-15-Fc fusion protein (hIL-15Rα-Fc) decreased...
the T-cell stimulatory and anti-apoptotic abilities of BXSB mouse macrophages in vitro [84]. Administration of the IL-15 blockade to affected mice reduced proteinuria and autoantibody titers in these mice [84].

IL-15 signaling has been inhibited in a clinical trial of rheumatoid arthritis using a recombinant IL-15 receptor-Fc fusion protein (HuMax-IL15/AMG 714) [89]. The trial was a phase I/II double-blind, placebo controlled dose-escalation study involving thirty patients. The drug was well-tolerated, with no treatment-related adverse effects. There is currently a clinical trial ongoing for this drug in treating psoriasis.

3.4. Interleukin 18

IL-18 is a proinflammatory cytokine that is primarily produced from antigen presenting cells such as macrophages and dendritic cells. IL-18 is also known as IFN-γ-inducing factor (IGIF) because it is involved in Th1 lymphocyte proliferation and promotes IFN-γ production. IFN-γ skewes the Th response towards a Th1 pattern, and promotes inflammation in lupus [90,91].

IL-18 is likely to be involved in several autoimmune disorders, including lupus. In lupus patients the level of IL-18 is elevated, especially in patients with lupus nephritis [91,92]. The inflamasome, which processes IL-1 and IL-18 to their active form, is also found at higher levels in lupus patients [93]. Mice that overexpress IL-18 develop lupus-like disease [94].

There are several potential strategies for inhibiting IL-18. Since IL-18 is a member of the IL-1 superfamily and it requires the intracellular cysteine protease caspase-1 for biological activity, inhibiting caspase-1 is one strategy to reduce IL-18 levels [95]. This method has been demonstrated to be effective in mice [95]. IL-18 can also be inhibited using neutralizing antibodies, which has also proven effective in mice [96]. An additional molecule shown to effectively inhibit IL-18 is IL-18 binding protein (IL-18BP), a natural inhibitor of IL-18. Along with the increased level of IL-18, IL-18BP is found at high levels in lupus nephritis with glomerular involvement. [49,97–100]. Although both IL-18 and IL-18BP are elevated in lupus, the balance favors IL-18, contributing to disease. Infusion of IL-18BP can greatly reduce INF-γ production, inhibiting the early Th1 response. This is effective in vitro and in mice [101–103].

In IL-18 deficient mice, NK cell activity and Th1 response was impaired [104]. Inhibition by anti-IL-18 antibodies induced a marked reduction in INF-γ and TNF-α and prevented experimental autoimmune encephalomyelitis [96]. Overexpression or introduction of IL-18BP ameliorates collagen-induced arthritis in murine models, as does neutralization with monoclonal antibodies [105,106].

Recombinant IL-18BP has been tested for safety and pharmacodynamics in rheumatoid arthritis and psoriasis in four studies [107], with good safety results. Safety studies are currently underway for an IL-18 neutralizing antibody (GSK1078086) in obese individuals. Targeting IL-18 is a promising therapy for reducing IFN-γ production and controlling the Th1 response, which is particularly important in the more severe forms of lupus, especially lupus nephritis.

4. Conclusions

With the advent of humanized monoclonal anti-cytokine antibodies, a new era of specific cytokine inhibition has emerged. These antibodies allow targeting of major pathogenic mechanisms in lupus, while generally sparing the immune system as a whole. The approval of Benlysta as a specific anti-lupus therapeutic has opened the doors for a new attack on multiple lupus pathogenic mechanisms. Given the complexity of lupus, it is likely that a multi-pronged approach, possibly with combinations of multiple cytokine inhibitors tailored to patients, will be necessary for these strategies to approach their promise in effectively treating lupus. Cytokine inhibition targets and mechanism will almost certainly be further refined in the future to allow not only for the wholesale inhibition of certain cytokines, but for the modulation and control of cytokine levels.

The use of cytokine inhibitors in lupus patients, while beneficial to the patients, also provides a window to further understand what goes wrong in lupus, what continues to cause damage, and which systems are dysregulated during the disease. This increased understanding will hopefully serve as a positive information feedback loop, allowing even better and more refined treatments for this difficult disease.

Conflict of interest statement

The author(s) declare that there are no conflicts of interest.

References


H.D. Verma, et al., Anti-nuclear antibody positivity and the use of certolizumab in inflammatory bowel disease patients who have had arthralgias or lupus-like reactions from infliximab or adalimumab, J. Dig. Dis. 12 (5) (2011) 379–383.


T. Araora et al., Differences in binding and effector functions between classes of TNF antagonists, Cytokine 45 (2) (2009) 124–131.


R. Kuhn, et al., Interleukin-10-deficient mice develop chronic enterocolitis, Cell 75 (2) (1993) 263–274.


C. Shimizu, et al., High circulating levels of interleukin-18 binding protein indicate the severity of glomerular involvement in systemic lupus erythematosus, Mod. Rheumatol. 22 (1) (2012) 73–79.


Appendix E. Review of interferon and apoptosis in systemic lupus erythematosus

See attached publication [243]. This work was included as Chapter 3 of a textbook about systemic lupus erythematosus. It is a review of current literature and summary. The majority of the introduction (Section 1) was taken from this work.
Interferon and Apoptosis in Systemic Lupus Erythematosus

Daniel N. Clark and Brian D. Poole
Brigham Young University
USA

1. Introduction
Systemic lupus erythematosus (SLE) is generally diagnosed long after the disease begins. This means that the cause of the disease is hard to find, buried in the past. In the search for the elusive causal agents for SLE, one candidate is the immune signaling molecule, interferon (IFN). Interferon is a secreted signaling protein, or cytokine, which is expressed at higher levels in SLE patients and has been associated with incidence and severity of the disease.

A combination of environmental triggers and genetic susceptibility combine to initiate SLE. Although there are many etiological components, they usually converge on a heightened state of activation for the immune system, with resultant increases in interferon production and interferon signaling. That is to say that interferon could be thought of as either a causative agent, a result of the disease, or both.

This chapter will discuss the basics of interferon function and how de-regulation of apoptosis can lead to interferon production due to immune complexes. We will then discuss how the functioning of the immune system changes in someone with SLE, the genes which are associated with risk for SLE, and clinical manifestations of interferon in SLE.

2. How interferon works in the context of SLE
Interferon is a signaling protein which is secreted to activate neighboring cells in response to viruses or other infections. It is a cytokine, or immune signaling molecule which allows communication between cells. When a cell is infected with a virus, interferon is produced and secreted as a warning to other cells to prepare for an infection. Interferons alpha (IFNα) and beta (IFNβ) are the type I interferons, and interferon gamma (IFNγ) is the type II interferon. Most of the cells in the human body have receptors for type I IFN, whereas certain immune cells express the receptor for type II IFN (Su, et al., 2004). The proteins are made by many different cells, but generally speaking, IFNα is of leukocyte origin, IFNβ is of fibroblast origin, and IFNγ is made by lymphocytes (Lucero, et al., 1982). Other less studied interferons also exist, and interferons are conserved among many species. This chapter will talk mostly about type I interferons, which are IFNα and IFNβ.

The main purpose of interferon is to shut down a cell before a virus can take it over, although it has many other jobs (Niewold, et al., 2010). Interferon signaling leads to increased apoptosis, which is a normal response to control viral spread or to decrease the
size of a tumor (Takaoka, et al., 2003). If one cell can undergo apoptosis before a virus can replicate and infect other cells, the infection is halted (Luker, et al., 2005).

Fig. 1. Interferon protein structures. Interferons alpha and beta, the type I interferons, have a common structure composed mainly of five alpha helices (shown are IFNa2a and IFNB1 based on PDB files 1itf and 1au1, respectively). Although the monomers of each are very similar in structure, the functional form of both is a dimer, and the two dimerize differently, IFNa2a along homologous surfaces and IFNB1 on opposing sides of the protein (Karpusas, et al., 1997). IFNg is shown in its dimerized form, with the two colors representing two intertwined monomers (based on PDB file 1hig). Not shown to scale; figures drawn with Jmol (Jmol, 2011).

Interferon can be produced in response to infection, other cytokines, mitogens and several signaling pathways. Once produced it is secreted where it can be recognized by other cells, which is called paracrine signaling, or by the cell which produced it, called autocrine signaling. One type of cell, the plasmacytoid dendritic cell (pDC) is a natural IFN producer, and is able to make very large amounts of IFNa (Ronnblom & Alm, 2001).

When interferon ligates an interferon receptor, signaling pathways are activated. Interferon causes an increase in the expression of both major histocompatibility complexes (MHC-I and MHC-II) for presentation of viral peptides to T cells, which can then lead to activation of other cells in order to kill infected cells, and remove them (Fruh & Yang, 1999). Interferon also increases intracellular levels of protein kinase R (PKR) which recognizes viral nucleic acids and activates RNase L to degrade viral RNAs. PKR also slows protein synthesis by inactivating translational initiation factors, so that viral proteins synthesis is slowed (Pindel & Sadler, 2011). p53 is also activated, which is pro-apoptotic (Takaoka, 2003). Interferons activate immune cells, especially natural killer cells and macrophages (Murray, 1988). This activation cascade is normally “turned off” after an infection is cleared to prevent damage to uninfected cells. However this activation state is not reduced to the normal levels in individuals with SLE, where a higher level of interferon is present (T. Kim, et al., 1987; Ytterberg & Schnitzer, 1982). This higher amount of interferon is also measurable by an increase in the expression of interferon-stimulated genes seen in lupus patients, called the interferon response signature (Baechler, et al., 2003; Bennett, et al., 2003; Feng, et al., 2006).
This means that IFN is turned on and that it is actively affecting how other cells are functioning.

![Diagram of cell signaling](image)

**Fig. 2.** Cell to cell IFN signaling and its effects. One cell produces interferon and either another cell (paracrine signaling) or the same cell (autocrine signaling) receives the signal. ↑: an increase, ↓: a decrease, MHC: major histocompatibility complex, PKR: protein kinase R, NK: natural killer cell, MΦ: macrophage

As a general feature of autoimmune diseases such as SLE, the immune system is in an “always on” state, which can lead to a breach in the body’s natural tolerance to self. Once this self tolerance is lost, autoimmune disease can result. In addressing why the immune system generates an attack against one’s own body, the over activation of the immune system, including the overproduction of interferon in SLE patients is a part of this picture.

### 3. Interferon leads to apoptosis, and the SLE-apoptosis connection

One effect of interferon production is the release of autoantigens due to increased cell death. This release is normally controlled by a process called efferocytosis, or apoptotic cell removal, where cell debris are processed by immune cells or neighboring cells which remove them by phagocytosis. Defects in apoptotic pathways have been noted in individuals with SLE (Gaipl, et al., 2006). Examples of why this occurs have been studied. For example, in SLE patients there is an overexpression of both soluble and membrane-bound Fas. Fas is a receptor which when ligated signals to a cell to undergo apoptosis. The levels of Fas also correlate with the amount of apoptotic lymphocytes and disease activity of SLE (Li, et al., 2009; Sahebari, et al., 2010). Mouse models of lupus commonly have genetic variations in apoptotic pathways such as the Fas/Fas L pathway and interferon pathways.

Mouse as well as human SLE patients make antibodies to self antigens. This is likely because of over-exposure of potential autoantigens to the immune system. This could be due to an increased amount of apoptosis, or a decrease in the rate of clearance of apoptotic debris. Apoptosis, which can be induced by interferon, is also part of the natural cycle of cellular growth and death. Cells undergoing apoptosis are recognized as dead by other cells, so that they are cleared (Munoz, et al., 2010).
Fig. 3. Production of interferon can begin with defects in apoptosis. This can be due to either an increase in apoptosis or a decrease in clearance of apoptotic debris. If contents are released, they can form immune complexes with autoantibodies. These immune complexes can cause cells to produce interferon. Manipulating this pathway is also a common characteristic of mouse models of lupus.

3.1 Mouse models allow the study of IFN and apoptosis pathways

Mouse models have been very useful in understanding the etiology and pathogenesis of lupus. Two approaches to experimental mice have been used to generate information about the role of interferon in lupus. In the first approach, interferon-related genes are knocked out and the resulting effects on lupus are studied. For the second, established lupus mouse models are studied on a molecular level for differences in interferon pathways or interferon-related effects. These two approaches often overlap, as in cases where interferon-related genes are knocked out in lupus-prone mice. Several established lupus mouse models include the MRL/lpr mice, NZW/NZB, and others. These are mice that spontaneously develop lupus, and several of them have been investigated to understand the role of interferon in their pathogenesis. Although a complete description of the mouse models for lupus is beyond the scope of this or any one publication, a few illustrative examples represent the power of these model systems.

One mouse model that is especially relevant for the study of interferon in lupus is the BXSB/MpJ (BXSB) or Yaa mouse. These mice spontaneously develop lupus-like disease in a sex-linked fashion because of a duplication of the Toll-like receptor 7 (TLR7) gene on the Y chromosome (Izui, et al., 1994). TLR7 is responsible for inducing interferon in response to viral infection or autoantibody production.

Another interesting mouse for the study of interferon is the NZB/NZW mouse. These mice spontaneously develop a lupus-like autoimmune disease. They have been used to investigate the role of several interferon-related molecules and cells. For example, treating these mice with interferon accelerates disease in a T-cell like manner (Z. Liu, et al., 2010; Mathian, et al., 2005), while knocking out or inhibiting interferon-related genes slows or eliminates the development of lupus-like symptoms (Jorgensen, et al., 2007; Sharma, et al., 2005). These mice have been used to clarify the interactions between sex hormones and...
interferon in lupus etiology (Bynote, et al., 2008; Panchanathan, et al., 2009; Panchanathan, et al., 2010), and they serve as an excellent all-around model for spontaneous development of lupus.

The role of several interferon-related molecules has been examined using a combination of mouse models. As an example, consider the gene interferon regulatory factor five (IRF5). This gene is an interferon-regulating gene which will be described in section 5.2 below. It was discovered that knockout of IRF5 prevents or inhibits the development of lupus in MRL/lpr mice, FcyR-/- Yaa mice, and pristine-injected mice (Richez, et al., 2010; Savitsky, et al., 2010; Tada, et al., 2011).

Mouse models for lupus represent a powerful and flexible mechanism for investigating the role of multiple aspects of lupus. However, it must be remembered that the mutations or disease manifestations in these mice are not necessarily related to those seen in human lupus, and therefore the results observed must be interpreted with caution.

4. Cycle of autoantibody production

When it comes to SLE we may think of interferon production as a cycle, which begins when an environmental trigger, such as a viral infection, UV light damage or medical treatment activates the immune system to produce interferon.

Normally B cells which produce antibodies to self-antigens undergo negative selection, where they receive signals to die off or become inactivated if they make antibody against a self-antigen. This self-tolerance is breached in SLE (Cancro, et al., 2009), and the self-antigens released from damaged or apoptotic cells during or after initial triggering events become the targets of autoantibodies. When autoantibodies are produced, they are made by B cells as well as plasma cells, which are a mature differentiated form of B cells.

Fig. 4. The altered immune response in SLE generates a cycle. In blue is a cycle which exists in SLE, amplifying the amount of IFN present. This cycle needs a trigger, but once it begins, it can leave the immune system in an “always on” state.
Autoantibodies lead to the production of interferon by forming immune complexes which are immunostimulatory (Ronnblom, et al., 2011). Immune complexes are composed of aggregates of antibody and antigen molecules which are processed by the body. These immune complexes are a main source of SLE pathology, as they obstruct small passages in areas of the body such as the kidneys and joints (Crispin, et al., 2010). Immune complexes may include the common SLE autoantigens such as RNA-containing protein complexes like Sm, RNP s, Ro, and La. Having a combination of both nucleic acids and protein complexed with antibody means many pathways can be turned on. For example, antibody can stimulate an immune cell through an Fc receptor, nucleic acids can stimulate cells through Toll-like receptors (TLRs), and proteins can be recognized by other antibodies.

Immune cells are activated by immune complexes and the cycle continues. Interferon production is instigated by immune cells which recognize part of the complex, be it the antibody, the antigen, or other associated molecules.

5. SLE genetic risk screens identify genes in interferon signaling pathways

We have looked at the disease state of SLE, and how the immune system functions improperly to instigate disease. Things begin when an environmental trigger works on the genetic background of varying degrees of susceptibility. Genetic susceptibility is thought to account for at least 20% of the risk for SLE (Deapen, et al., 1992). To find the actual genes involved, studies are performed to determine the linkage or association of a variation in the genome to a particular disease.

One important method is called a genome wide association study (GWAS). These GWA studies genotype thousands of individuals, grouped into SLE patients and non-patients comparing them at thousands of single nucleotide polymorphisms (SNPs). These studies reveal the genomic regions which contain disease-associated genes, because the variations are more common in people with the disease. Individual genes or gene pathways are pinpointed, and can ultimately lead to treatment strategies. Many genes have been identified that contain SNPs which confer risk to SLE.

These studies are especially useful for diseases with unknown or complex genetic components. The genome is examined for sets of single nucleotide polymorphisms (SNPs). When sets of SNPs are usually inherited together in a group it is called a haplotype. When a haplotype is more common in the disease group than in the unaffected group, it can be assumed that it is associated with the disease. Although specific genes are sometimes found which may predict a disease, it is more likely that the information will reveal molecular pathways associated with the disease. Association of genes or pathways to diseases such as heart disease, asthma, diabetes and others have been found using this method (Stranger, et al., 2011). The amount of effect is measured as an odds ratio (OR), which is a measure of the strength of association of the disease with a haplotype. A median OR value is around 1.3, with some genes having much higher association ORs. For example, one of the lupus-associated haplotypes TREX1, has a published OR of 25 (Lee-Kirsch, et al., 2007). In such cases, the genetic risk is almost certainly associated with the disease.

An important caveat to these tests is that they answer the question, “What?” but not the question, “How?” That is, they identify genetic loci which confer risk to SLE, but then further studies are needed to show what functional changes affect people with a risk haplotype. For most of the genes, we do not know what functional role they play. However
it is promising to note that the genes are within certain pathways, some of which are already associated with lupus.

Fig. 5. Genome wide association studies (GWAS). Genome-wide association studies aim to discover the genetic risk component of a disease by finding differences in a group with the disease compared to a group of unaffected control individuals.

Several review articles have reviewed the findings of many lupus GWAS with varying degrees of certainty (R.R. Graham, et al., 2009; I.T.W. Harley, et al., 2009; Moser, et al., 2009; Rhodes & Vyse, 2008; Sebastiani & Galeazzi, 2009). In some cases the indicated susceptibility genes are common in many ethnicities and populations, while others are specific to certain groups. The statistical significance of many of these genes is well established, while others are novel and need to be replicated by other groups. An important finding is that most of the genes that have been identified in GWA studies can be grouped into several functional pathways. We will focus on the genes in the IFN pathway and the pathways involving clearing of apoptotic cells and immune complexes.

5.1 Interferon production pathways
Intracellular signaling pathways which control interferon production include the production of type I interferons by interferon regulatory factors (IRFs), and the production of type II interferon by STAT4. IRFs are activated by TLRs, which are extracellular or endosomal pattern recognition molecules. TLRs 7, 8, and 9 recognize nucleic acids and are endosomal. Maintaining these TLRs in the endosome instead of the cell surface is an important barrier to too frequent TLR activation. Once the nucleic acids are brought into the cells through endocytosis, the TLRs become activated to turn on IRFs. TLR 8 and TLR 9 have both been identified as lupus risk genes (Armstrong, et al., 2009; Xu, et al., 2009).

TLRs begin a signaling cascade through a MyD88 signaling complex. MyD88 activates another confirmed locus of SLE risk, the gene which encodes IL1 receptor-associated kinase 1
In Sle1 and Sle3 mouse models of lupus, IRAK deficiency eliminated most lupus symptoms (Jacob, et al., 2009), which highlights the importance of IRAK1. Since this gene is on the X chromosome, it could help explain why lupus is more common among women. The MyD88 complex can be affected by osteopontin (OPN). It regulates IFNα production in plasmacytoid dendritic cells, which are the body’s main IFNα producer cell (Cao & Liu, 2006). The lupus-risk variant of OPN was tied to high IFNα levels in certain lupus patients (Kariuki, et al., 2009b).

Two interacting proteins involved in inflammation, TNFα-induced protein 3 (TNFAIP3) and TNFAIP3-interacting protein 1 (TNIP1), are also lupus risk loci (Gateva, et al., 2009; Musone, et al., 2008). TNFAIP3 encodes the protein A20, which abrogates NFκB after an inflammatory response, and lupus-risk variants of this gene are associated with blood and kidney manifestations (Bates, et al., 2009). TNIP1 interacts with TNFAIP3 as well as affecting several other signal transduction pathways.

Interferon regulatory factors are activated next, downstream of TLRs; they are transcription factors which travel to the nucleus to bind DNA to initiate transcription. IRF5 binds to a sequence-specific region of DNA to induce IFN production. It has been confirmed as a risk factor for SLE in among several ethnicities (Kawasaki, et al., 2008; Kelly, et al., 2008; Lee & Song, 2009; Reddy, et al., 2007; Shimane, et al., 2009). There are three main genetic variants within IRF5, one copy number variant with either two or four copies of a 30-bp sequence, and two SNPs (R.R. Graham, et al., 2007b). The rs2004640 SNP changes the first exon, although this exon does not encode protein. The other SNP, rs10954213, creates an early polyadenylation sequence, which yields shorter more stable mRNA (D.S.C. Graham, et al., 2007a). Work has shown that these variants increase the amount of IFN in the presence of SLE autoantibodies (Niewold, et al., 2008; Salloum, et al., 2009).

![Interferon production pathways](image)

Fig. 6. Interferon production pathways are affected by lupus-risk genes. The * represents genes which have been identified as having risk for lupus. The endosomal TLRs (7, 8, and 9) can bind to autoantigenic nucleic acids and signal through a MyD88 complex which can be affected by association with osteopontin (OPN). If it is not blocked by TNFAIP3, this activates an IRAK signaling complex to phosphorylate the IRF5 and IRF7 transcription factors to produce type I IFN. IL-12 or IL-23 signal through Tyk2/Jak2 to activate the STAT4 transcription factor to produce type II IFN, commonly in T helper cells (Watford, et al., 2004).
IRF7 is associated with SLE risk by its proximity to SNPs in the IRF7/KIAA1542 locus (J.B. Harley, et al., 2008; Suarez-Gestal, et al., 2009). IRF7 SNPs have been shown to lead to increased IFNα levels and alter of which autoantibodies are made (Salloum, et al., 2009). Signal transducer and activator of transcription 4 (STAT4) is also associated with risk for SLE. It is a transcription factor which activates genes in proliferation, differentiation and apoptosis pathways. Two STAT4 SNPs have been examined, rs7574865 increases sensitivity to IFNα (Kariuki, et al., 2009a), and rs3821236 causes STAT4 to be transcribed at higher levels and is additive with IRF5 risk loci so that when both are present, the risk to SLE is multiplied (Abelson, et al., 2009; Sigurdsson, et al., 2008).

5.2 Genes associated with apoptosis and immune complexes
Another set of risk genes can be placed into a functional group of apoptosis-associated genes. As we read earlier in the chapter, defects in apoptosis can lead to the presence of potential autoantigens. For example, a cell undergoes apoptosis and instead of being cleared by other cells, its contents are released. The cellular contents can contain things like nucleic acids, RNA binding proteins, and others which are common lupus autoantigens. If antibodies bind to these antigens, a complex of multiple antibodies and multiple antigens can aggregate. The resultant immune complexes can be broken down through reactions with complement components, which are commonly found at low levels in SLE patients (C.C. Liu & Ahearn, 2009). If they are not broken down, they reach areas such as the kidneys or joints, which can be damaged by these immune complexes. This is how organ damage usually occurs in lupus patients.

Fig. 7. Genes associated with risk for lupus in the apoptosis pathway. The * represents genes which have been identified as having risk for lupus. TNFα, CASP10 and IRF5 are pro-apoptotic whereas OPN and p21 are anti-apoptotic. These genes all have a role in how much apoptosis is occurring. Once apoptosis has transpired, the cell must be cleared. Parts of apoptotic cells or immune complexes can be recognized by other cells to facilitate their removal. This is aided by recognition molecules such as the complement components shown here.

The problem of creating autoantibodies could stem from too much apoptosis or too little clearance of apoptotic debris. Genes identified in GWA studies that could alter the amount of apoptosis include TNFα, caspase 10, IRF5, osteopontin and p21.

TNFα was identified as a risk factor for lupus in certain ethnicities (Jimenez-Morales, et al., 2009; Lin, et al., 2009). TNFα is a cytokine which is produced and secreted to signal to
other cells and is found at high levels in the serum of lupus patients (Davas, et al., 1999; Emilie, et al., 1996; Sabry, et al., 2006). Part of its function is to induce apoptosis—when a cell binds TNFα, it activates the caspase cascade. Caspases are proteases which are activated under certain conditions and are a hallmark of apoptosis. They cleave other caspases as well, and the combined proteolytic activity of several different activated caspases breaks down cellular components as the cell prepares to die. Caspase 10 is part of this cascade and is another lupus susceptibility gene (Armstrong, et al., 2009). Caspase 8, is activated by TNF signaling, and cleaves caspase 10, which then cleaves caspases 3 and 7. IRF5, as well as being a transcription factor which helps produce IFN, is also a tumor suppressor gene which is commonly inactivated in cancers. This is because of IRF5’s pro-apoptotic function.

Osteopontin (OPN) and p21 are also lupus risk genes, both anti-apoptotic. OPN promotes proliferation, as well as prevention of death under apoptotic stimuli (Standal, et al., 2004). A mimic of p21 was used in the treatment of murine lupus in the NZB/NZW mouse, and it was found to dramatically reduce the disease (Gouvelstre, et al., 2005).

So, there are genes which dysregulate the amount of apoptosis, and they are associated with risk for lupus. But this is only half of the picture; the other part is the clearance of apoptotic cells or immune complexes. Several SLE susceptibility genes in this pathway have been identified as well. Active SLE can be assessed when low levels of complement proteins are found in circulation. Complement can function against microbes during an infection, but can also help to degrade immune complexes. Once attached, they can help cells recognize and degrade them. Other proteins function to bind apoptotic cells or immune complexes to facilitate their uptake by other cells.

Integrin αM (ITGAM) has been convincingly associated to SLE (Nath, et al., 2008). Risk variants of ITGAM have been associated with certain clinical manifestations of lupus (Kim-Howard, et al., 2010). It is a cell receptor which binds to OPN or to complement C3b. C3b binds to apoptotic cells or immune complexes.

SLE association with complement components C1q, C2, C4a and C4b have large OR values, meaning that the risk haplotypes of these genes are causing a large effect. When C1q is expressed at low levels it can lead to lupus, and it was shown to increase the amount of IFN produced due to immune complexes (Lood, et al., 2009). Complement components function by binding immune complexes by the Fc region of antibody or by binding to other parts of apoptotic cells, which can opsonize them for easier uptake by other cells. Cells can then remove the immune complex or apoptotic debris by endocytosis. Receptors for the Fc region of antibody have also been implicated in SLE risk (Lee-Kirsch, et al., 2007). These receptors can bind to antibody within an immune complex.

Other proteins such as milk fat globule EGF factor 8 (MFG-E8) and C-reactive protein (CRP) can bind to apoptotic cells by recognizing phospholipids on their membranes. MFG-E8 binds to phosphatidylserine, an “eat me” signal which is expressed on apoptotic cells. The MFG-E8 knockout mouse gets SLE because of failure to remove apoptotic cells (Yamaguchi, et al., 2010). CRP binds to phosphocholine, which is present on dying or damaged cells. Both MFG-E8 and CRP are lupus risk genes (Batuca & Alves, 2009; Hu, et al., 2009; H.A. Kim, et al., 2009). Low mannose-binding lectin (MBL) levels can lead to higher levels of apoptosis is this lupus-risk associated gene (Pradhan, et al., 2010).
The number of genes associated with risk for SLE will likely increase, though we have an interesting pool of genes already that point to certain pathways associated with the disease. The interferon and apoptosis pathways are certainly important in SLE etiopathogenesis.

6. Clinical component of interferon and SLE

Many researchers have sought to determine if higher levels of IFN, which is common in lupus patients, is a cause of lupus or an effect of lupus. An interesting occurrence can happen when someone undergoes treatment with IFNα. The presence of increased levels of IFN leads to lupus or a lupus-like syndrome (Gota & Calabrese, 2003; Ioannou & Isenberg, 2000; Niewold & Swedler, 2005). Because the lupus symptoms usually disappear after IFN treatment ends, this connection suggests that IFN may be more of a cause than an effect. In a small number of cases, some patients also develop SLE as a result of these IFN treatments. Furthermore, within a family, the levels of interferon among all members correlate, suggesting that this is a heritable trait (Niewold, et al., 2007). That is, even the siblings of a lupus patient with high IFN levels are more likely to have higher IFN levels. This also supports a causal role for IFN.

Clinically, disease activity can be measured and correlated to other observations to determine the cause of the different levels of activity. One item linked to SLE activity is interferon, where higher levels of IFN in the serum correlated with more severe disease in most cases (Bauer, et al., 2009; Dall’Era, et al., 2005; Feng, et al., 2006; Landolt-Marticorena, et al., 2009; Petri, et al., 2009; Zhuang, et al., 2005).

Common autoantibodies also correlate with IFN levels. A very strong correlation is consistently observed between IFNα levels and the presence of antibodies to common SLE autoantigens like Ro, La, Sm, RNP, and dsDNA (Kirou, et al., 2005).

Another set of findings has to do with properties of main producer of IFNα, the plasmacytoid dendritic cells (pDCs). High numbers of IFN-producing pDCs have been observed in lupus skin lesions (Blomberg, et al., 2001; Farkas, et al., 2001). Since the cells are present at the scene of the crime, the increased interferon could have to do with the pathology in these cases.

At the time of writing, two clinical drug trials for SLE are being conducted, Sifalimumab is in Phase II, and Rontalizumab is in Phase I. Both are antibodies, designed to block interferon alpha signaling by binding it to prevent its recognition by neighboring cells (Clinical Trials, 2011). If these drugs are found to be effective, it will show that IFN plays a critical role in the pathogenesis of lupus. In addition, the United States Food and Drug Administration recently approved an antibody to B lymphocyte stimulator (BLyS) to treat SLE called Belimumab (Sanz, et al., 2011). This should help control the selective apoptosis and autoantibody production to some degree.

7. Conclusions

Several themes have been examined in this chapter. Specifically that the production of interferon is tied to lupus and that apoptosis, clearance of apoptotic cells, and the formation of immune complexes are events that can augment the production of interferon. Exciting findings about the actual genetic causes of SLE are being examined which will lead to better treatments for this complex disease. Although most of the data discussed in this chapter are inferential, there is a large body of evidence in support of the hypothesis that increased interferon signaling promotes an autoimmune state in those genetically prone to SLE.
8. References


Dall’Era, M.C.; Cardarelli, P.M.; Preston, B.T.; Witte, A. & Davis, J.C. (2005). Type I interferon correlates with serological and clinical manifestations of SLE. Animals of the Rheumatic Diseases, 64 pp. 1692-1697


alpha is a common genetic risk factor for asthma, juvenile rheumatoid arthritis, and systemic lupus erythematosus in a Mexican pediatric population. *Human Immunology*, 70 pp. 251-256

Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org


Niewold, T.B.; Clark, D.N.; Salloum, R. & Poole, B.D. Interferon Alpha in Systemic Lupus Erythematosus. *Journal of Biomedicine and Biotechnology*, pp. 8


Panchanathan, R.; Shen, H.; Zhang, X.; Ho, S.M. & Choubey, D. (2010). Mutually positive regulatory feedback loop between interferons and estrogen receptor-
alpha in mice: implications for sex bias in autoimmunity. *PLoS ONE*, 5 pp. e10868


Ronnblom, L.; Alm, G.V. & Eloranta, M.L. The type I interferon system in the development of lupus. *Seminars in Immunology*, 23 pp. 113-121


Sahebari, M.; Hatef, M.R.; Rezaieyazdi, Z.; Abbasi, M.; Abbasi, B. & Mahmoudi, M. Correlation between Serum Levels of Soluble Fas (CD95/Apo-1) with Disease Activity in Systemic Lupus Erythematosus Patients in Khorasan, Iran. *Archives of Iranian Medicine*, 13 pp. 135-142


131


Appendix F. Microarray analysis of B cell expression differences due to an rs2004640 risk haplotype

See attached publication [192]. This work was performed in collaboration with the Oklahoma Medical Research Foundation under Dr. Brian Poole. Parts of the results, discussion and materials and methods were included in this dissertation (Sections 2.6, 3.4, and 4.11).
Research Article

Effects of IRF5 Lupus Risk Haplotype on Pathways Predicted to Influence B Cell Functions

Joel M. Guthridge, 1 Daniel N. Clark, 2 Amanda Templeton, 1 Nicolas Dominguez, 1 Rufei Lu, 1 Gabriel S. Vidal, 1 Jennifer A. Kelly, 1 Kenneth M. Kauffman, 3 John B. Harley, 3 Patrick M. Gaffney, 1 Judith A. James, 1, 4 and Brian D. Poole 1, 2

1 Arthritis and Clinical Immunology Program, Oklahoma Medical Research Foundation, 825 NE 13th Street, Oklahoma City, OK 73104, USA
2 Department of Microbiology and Molecular Biology, Brigham Young University, 857 WIDB, Provo, UT 84604, USA
3 Division of Rheumatology, Cincinnati Children's Hospital, 3333 Burnet Avenue, Cincinnati, OH 45229, USA
4 Departments of Medicine and Pathology, The University of Oklahoma Health Sciences Center, 1100 N. Lindsay, Oklahoma City, OK 73104, USA

Correspondence should be addressed to Brian D. Poole, bpoole@gmail.com

Received 5 May 2011; Revised 4 November 2011; Accepted 5 November 2011

Academic Editor: Timothy B. Niewold

Copyright © 2012 Joel M. Guthridge et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Both genetic and environmental interactions affect systemic lupus erythematosus (SLE) development and pathogenesis. One known genetic factor associated with lupus is a haplotype of the interferon regulatory factor 5 (IRF5) gene. Analysis of global gene expression microarray data using gene set enrichment analysis identified multiple interferon- and inflammation-related gene sets significantly overrepresented in cells with the risk haplotype. Pathway analysis using expressed genes from the significant gene sets impacted by the IRF5 risk haplotype confirmed significant correlation with the interferon pathway, Toll-like receptor pathway, and the B-cell receptor pathway. SLE patients with the IRF5 risk haplotype have a heightened interferon signature, even in an unstimulated state ($P = 0.011$), while patients with the IRF5 protective haplotype have a B cell interferon signature similar to that of controls. These results identify multiple genes in functionally significant pathways which are affected by IRF5 genotype. They also establish the IRF5 risk haplotype as a key determinant of not only the interferon response, but also other B-cell pathways involved in SLE.

1. Introduction

Systemic lupus erythematosus is a complex disease with multifactorial etiology and pathogenesis. Studies in identical twins indicate that concordance for lupus is approximately 40%, indicating a strong but not exclusive genetic component [1, 2]. Recent genetic analyses have identified more than thirty candidate genes that are associated with lupus risk [3–18]. IRF5 was found to be associated with lupus by multiple independent groups in a variety of populations [10, 13–15, 19, 20]. IRF5 risk haplotypes may function at the crossroads of environmental risk, such as virus infection, and cellular immune responses. At least three polymorphisms of IRF5 have been identified that contribute independently to the risk for lupus, which together constitute the lupus risk haplotype [10, 21]. Although the majority of the polymorphisms that have been associated with lupus are in nontranslated regions, they may affect several facets of IRF5 activity, including splicing, RNA stability, transcription factor binding, and apoptosis [9, 10, 15, 21, 22].

IRF5 is important in the production of and response to interferon alpha (IFNα), which is heightened in lupus. IFNα is produced by dendritic cells, macrophages, B cells, and other cell types, primarily in response to virus infection [23, 24]. Dendritic cells have been shown to produce IFNα in response to incubation with immune complex-containing...
sera from SLE patients [25], especially patients that have the risk haplotype for IRF5 [19]. Additionally, serum interferon levels, as well as the interferon response signature, are increased in patients with the risk haplotype [19, 26]. IRF5 is an especially interesting candidate for a genetic risk factor in lupus because it acts in pathways that control many of the cellular and immune responses to environmental factors, such as infection, which may contribute to lupus.

One putative environmental agent that is strongly associated with risk for lupus is Epstein-Barr virus (EBV) infection. Lupus has been associated with prior EBV infection in both pediatric and adult populations [27–33]. EBV expresses antigens that are immunologically cross-reactive with significant lupus autoantigens such as Sm and nRNP [34–39]. However, since over 95% of adults are infected with EBV, determining why EBV could contribute to lupus in certain individuals but not others has proven challenging. The identification of IRF5 and other genetic risk factors for lupus open the possibility that the lupus-associated genetic polymorphisms in one or more of these genes works in concert with environmental factors culminating in the increased observed risk for developing lupus. Previous work has shown that pediatric lupus patients have broadened, more cross-reactive humoral immune responses to EBV than controls [40]. EBV is also not as well controlled in lupus because it acts in pathways that control many of the cellular and immune responses to environmental factors, such as viral infection or the response to viral infection conferred by genetic factors such as IRF5 polymorphisms may in part explain these observations.

Since B cells are the primary host cell for EBV infection, we used B cells and EBV interactions as a model to study the impact of IRF5 genotype on downstream B-cell responses. For this study, we examined differences in B-cell gene expression between naïve B cells from individuals with the IRF5 risk haplotype and those with the protective or neutral haplotypes at both basal levels and after exposure to EBV. Naïve B cells were chosen because they are the cell type in which EBV establishes latent infection [43]. We found multiple networks of genes that were enriched for differential expression, as well as individual gene expression differences. Most importantly, we identified different expression patterns of interferon response genes in lupus patients based on the IRF5 risk haplotype. Understanding these differences will aid in determining mechanisms through which the genetic risk conferred by the IRF5 risk haplotype is manifested.

2. Materials and Methods

2.1. Study Participants. Genotypes were previously collected on samples obtained from the Oklahoma Rheumatic Disease Resource Cores Center (ORDRCC) at the Oklahoma Medical Research Foundation. Previously enrolled subjects were contacted for study participation based upon their IRF5 risk and protective haplotypes using genotypes at single nucleotide polymorphisms rs2004640 and rs10954213. Five IRF5 high-risk (3 controls, 2 patients) and five IRF5 nonrisk (2 controls, 3 patients) sex- and race- matched individuals were recruited. The study was approved by the institutional review board at OMRF and OUHSC, and informed consent was obtained from all subjects in the study.

2.2. B-Cell Stimulation. Peripheral blood mononuclear cells were separated by density gradient centrifugation from the peripheral blood of volunteers. Naïve B cells were isolated using the MACS Naïve B Cell Isolation Kit II (Miltenyi Biotec Inc). Untouched naïve B cells were incubated at a 1:1 (v/v) ratio with either virus-free media or infectious EBV for 16 hours. Virus preparations were in the form of B95-8 cell culture supernatant. The same preparation of supernatant was used for all assays.

2.3. Gene Expression Profiling. Total cellular mRNA was purified from lysates of infected and mock-infected cells using the Ambion RNaqueous-Micro Kit (Applied Biosystems, Austin, TX, USA) according to the manufacturer’s protocol and quantified using a NanoDrop spectrophotometer (Nanodrop Technologies, Inc.). cRNA amplification and labeling with biotin were performed using the Illumina TotalPrep RNA amplification kit protocol (Ambion, Austin, TX, USA) on an aliquot of 200 ng of total RNA. Whole genome expression analysis was performed using the Illumina HumanRef-8 v3 gene expression chip (24,526 transcripts) following the Illumina Whole-Genome Expression Protocol.

2.4. Statistical and Pathway Analysis. The microarray data were analyzed using gene set enrichment analysis, and pathway analysis to investigate changes in gene networks. These analyses were followed by comparison of individual gene expression differences inside these networks. Raw expression data was first normalized using the MDAT toolbox [44]. Gene Set Enrichment Analysis software (Molecular Signatures Database) was used to determine whether an a priori functionally defined set of genes showed statistically significant, concordant differences between two phenotypes (IRF5 risk and nonrisk haplotypes) [45, 46]. Significant gene sets were identified by an enrichment score, which reflects the degree to which a gene set is overrepresented at the top or bottom of a ranked list of genes, and a false discovery rate (FDR) of <25%. We focused our subsequent pathway analysis on the subset of enriched genes (n = 368) from the statistically significant gene sets.

Pathway analyses were generated through the use of Ingenuity Pathways Analysis [47] (Ingenuity Systems, http://www.ingenuity.com/). A data set containing gene identifiers and corresponding expression values was uploaded into in the application. Each gene identifier was mapped to its corresponding gene object in the Ingenuity Pathways Knowledge Base. The expression values entered were the normalized log (intensity) values of IRF5 nonrisk and risk haplotype individuals, respectively.

Canonical pathways analysis identified the pathways from the Ingenuity Pathways Analysis library of canonical pathways that were most significant to the data set. The significance of the association between the data set and the canonical pathway was measured in two ways: (1) a
ratio of the number of genes from the data set that map to the pathway divided by the total number of genes that map to the canonical pathway; (2) Fisher’s exact test was used to calculate a P value determining the probability that the association between the genes in the dataset and the canonical pathway is explained by chance alone. All associations are supported by at least one reference from the literature, from a textbook, or from canonical information stored in the Ingenuity Pathways Knowledge Base. Human and mouse orthologs of a gene are stored as separate objects in the Ingenuity Pathways Knowledge Base, but are represented as a single node in the network. Heat maps were created with previously mentioned expression values in Spotfire software.

3. Results

3.1. Enrichment of Lupus-Related Gene Sets by IRF5 Risk Haplotype. Using existing genotyping data for single-nucleotide polymorphism rs2004640 (T: risk, G: protective) and rs10954213 (A: risk, G: protective) [10, 11, 20, 21, 47] the frequency of the risk, protective, and neutral IRF5 haplotypes were examined in 1,390 SLE patients and 2,039 controls enrolled in a large cohort of European American female (EA) SLE patients from the Lupus Family Registry and Repository (LFRR: http://lupus.omrf.org/). A significant enrichment for the risk and risk-neutral haplotypes was found in lupus patients (P < 0.0001), as expected (Table 1). Using these results, five individuals with the risk or risk-neutral haplotypes and five with the protective or protective-neutral haplotypes were recruited from the local SLE collections through the ORDRCC for further study. These included two SLE patients and three controls in the risk groups, and three SLE patients and two controls with protective and protective-neutral haplotypes. The risk group was enriched for controls so that the effects of the IRF5 haplotypes could be better studied in the absence of potential confounding genetic influences or factors related to lupus.

In all, 368 genes were found to be significantly (P < 0.05) differentially expressed between the risk and the nonrisk individuals. Gene set enrichment analysis was used to look for gene pathways overrepresented when comparing expression data from the risk haplotype and the protective haplotype phenotypic groups in the case or control groups separately. This analysis examines 6,769 a priori defined functional gene sets [46, 48]. In the unaffected controls with the risk haplotype, nine gene sets were enriched with a false discovery rate (FDR) <25% (Table 2), and 19 were enriched with a nominal P value <0.01, but an FDR >25% (Supplemental Table 2 which available online at doi:10.1155/2011/594056). In the lupus patients with the risk haplotype five gene sets were enriched with an FDR <25%. However, four of these were different versions of interferon-alpha gene sets. The fifth was the lupus-related interferon response signature (Table 2). Seven gene sets were enriched with a nominal P value <0.01 but a FDR >25% in the SLE risk haplotype cells, including the Toll-like receptor (TLR) gene set (Supplemental Table 1).

Fewer gene sets were as highly enriched in the protective haplotype cells. No gene sets were significantly enriched at an FDR level <25% in either the unaffected controls or the lupus patients with the protective haplotype. However, 39 gene sets were enriched at P < 0.01 in the SLE patients with the protective haplotype, and 35 gene sets were enriched at P < 0.01 level in the controls with the protective haplotype (Supplemental Tables 3 and 4).

3.2. Identification of Lupus-Related Pathways Differentially Affected by IRF5 Haplotype. Genes from the enriched gene sets described above which also demonstrated differences in expression in the previous analysis were included in a global pathway analysis using the Ingenuity Pathway Analysis system. This analysis uses the curated Ingenuity Knowledge Base to associate sets of genes and expression data with established gene pathways. Fisher’s exact test was used to quantify the degree of association with these pathways. When the cells with the risk haplotype (both stimulated and unstimulated) were compared to those with the non-risk haplotypes, three canonical pathways were found with statistically significant (P < 0.01) association: the interferon (Figure 1), Toll-like receptor (Figure 2), and B-cell receptor (Figure 3) pathways. Interestingly, all three of these pathways have significant implications for lupus. These three canonical pathways and relative changes in expression following EBV exposure are represented in Figures 1–3 and Table 3.

3.3. Identification of Individual Genes Differentially Expressed by IRF5 Haplotype. Several genes in the interferon pathway exhibited differential expression between either the risk and protective haplotypes or the EBV exposed and unexposed conditions. Genes with differential expression were selected based on inclusion in a significantly associated pathway, average expression values of at least thirty for one condition, and differential expression of at least 1.5-fold. Differential expression comparisons were done both with the unstimulated and the EBV-infected states (Table 3). Several genes were differentially expressed in the interferon pathway, including interferon-induced transmembrane protein 1 (IFITM1), signal transducer and activator of transcription 1 (STAT1), IFNα receptor 2 (IFNAR2), 2′–5′-oligoadenylate synthetase 1 (OAS1), and MX1. The expression patterns of these genes varied based on IRF5 haplotype and EBV infection status (Figure 1).

IFITM1 was more strongly expressed in the risk cells than in the nonrisk in the unstimulated condition (2.1-fold difference). When the cells were exposed to EBV, this difference disappeared, a result of a greater increase in expression (4.8-fold) in the nonrisk cells than the risk (2.8-fold). STAT1 was slightly underexpressed in the risk cells than in the nonrisk prior to EBV exposure (1.3-fold), but after EBV exposure it was more highly expressed in the risk cells (1.3-fold). IFNAR2 is an interferon receptor that contributes directly to the response to interferon, making this
Table 1: Frequency of IRF5 haplotypes in lupus patients and healthy controls. Haplotype frequencies observed in controls and systemic lupus erythematosus (SLE) patient cohort at single-nucleotide polymorphisms rs2004640 (T: risk, G: protective) and rs10954213 (A: risk, G: protective).

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Allele 1</th>
<th>Allele 2</th>
<th>Frequency (patients)</th>
<th>Frequency (controls)</th>
<th>SLE risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TA</td>
<td>TA</td>
<td>0.2576</td>
<td>0.1947</td>
<td>Risk/risk</td>
</tr>
<tr>
<td>2</td>
<td>TA</td>
<td>TG</td>
<td>0.4460</td>
<td>&lt;0.0001</td>
<td>Risk/neutral</td>
</tr>
<tr>
<td>3</td>
<td>GA</td>
<td>GG</td>
<td>0.2108</td>
<td>0.2737</td>
<td>Protective/protective</td>
</tr>
<tr>
<td>4</td>
<td>GG</td>
<td>TG</td>
<td>0.0165</td>
<td>0.1810</td>
<td>Protective/neutral</td>
</tr>
</tbody>
</table>

Table 2: Effect of the IRF5 risk haplotype on the expression of gene sets. Gene set enrichment analysis showed gene sets enriched in the risk haplotypes of either SLE-unaffected controls or SLE patients without EBV infection. Genes shown have a false discovery rate (FDR) of <25%.

<table>
<thead>
<tr>
<th>Gene set name</th>
<th>Affected pathways or cellular conditions</th>
<th>No. of genes</th>
<th>Normalized enrichment score</th>
<th>P value</th>
<th>FDR q-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unaffected controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CROONQUIST_IL6_STROMA_UP</td>
<td>IL-6 exposure</td>
<td>37</td>
<td>-2.026</td>
<td>0.0018</td>
<td>0.038</td>
</tr>
<tr>
<td>PASSERINI_INFLAMMATION</td>
<td>Inflammation</td>
<td>23</td>
<td>-1.751</td>
<td>0.0112</td>
<td>0.212</td>
</tr>
<tr>
<td>PASSERINI_PROLIFERATION</td>
<td>Proliferation</td>
<td>62</td>
<td>-1.782</td>
<td>&lt;0.0001</td>
<td>0.224</td>
</tr>
<tr>
<td>ADIP_DIFF_CLUSTER2</td>
<td>Differentiation</td>
<td>41</td>
<td>-1.752</td>
<td>0.004</td>
<td>0.225</td>
</tr>
<tr>
<td>CROONQUIST_RAS_STROMA_DN</td>
<td>Ras activation</td>
<td>21</td>
<td>-1.757</td>
<td>0.0038</td>
<td>0.229</td>
</tr>
<tr>
<td>UVB_NHEK3_C6</td>
<td>UV light exposure</td>
<td>27</td>
<td>-1.763</td>
<td>0.0039</td>
<td>0.234</td>
</tr>
<tr>
<td>HOHENKIRK_MONOCYTE_DEND_DN</td>
<td>Dendritic cell maturation</td>
<td>121</td>
<td>-1.724</td>
<td>0.0348</td>
<td>0.239</td>
</tr>
<tr>
<td>LEE_DENA_UP</td>
<td>Murine liver cancer</td>
<td>59</td>
<td>-1.769</td>
<td>&lt;0.0001</td>
<td>0.241</td>
</tr>
<tr>
<td>ZUCCHI_EPITHELIAL_DN</td>
<td>Breast cancer metastasis</td>
<td>44</td>
<td>-1.785</td>
<td>0.0117</td>
<td>0.245</td>
</tr>
<tr>
<td>SLE patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IFNALPHA_HCC_UP</td>
<td>IFNα</td>
<td>29</td>
<td>-1.968</td>
<td>0.0038</td>
<td>0.042</td>
</tr>
<tr>
<td>IFNALPHA_NL_HCC_UP</td>
<td>IFNα</td>
<td>18</td>
<td>-1.878</td>
<td>0.0096</td>
<td>0.063</td>
</tr>
<tr>
<td>RADAeva_IFNA_UP</td>
<td>IFNα</td>
<td>49</td>
<td>-1.897</td>
<td>0.0059</td>
<td>0.07</td>
</tr>
<tr>
<td>IFNALPHA_NL_UP</td>
<td>IFNα</td>
<td>27</td>
<td>-1.847</td>
<td>0.0099</td>
<td>0.075</td>
</tr>
<tr>
<td>BENNETT_SLE_UP</td>
<td>SLE</td>
<td>28</td>
<td>-1.785</td>
<td>0.0082</td>
<td>0.138</td>
</tr>
</tbody>
</table>

Gene of interest that are overexpressed in the risk cells in the TLR pathway without EBV exposure include CD14 (3.2-fold), lymphocyte antigen 96 (LY96, or MD-2) (2.3-fold), and TLR1 (1.7-fold).

The B-cell receptor (BCR) pathway exhibited differences in gene expression due to the IRF5 risk haplotype (Figure 3). CD79A and CD79B, which together form part of the BCR, were both downregulated 2.2-fold in the uninfected risk cells, but this difference disappeared after EBV infection. Ras-related C3 botulinum toxin substrate 1 (RAC1), a gene involved in lymphocyte differentiation and survival [49], was overexpressed in risk cells under all conditions (1.5-fold in mock infected cells, 1.7-fold in EBV infected cells). Expression of the signaling protein AKT1 and the transcription factor NFκB2 were downregulated in EBV infected cells with the IRF5 risk haplotype (1.9- and 3.4-fold, resp.). In three other genes, phosphatidylinositol 3 kinase catalytic subunit α (PIK3CA), nuclear factor of activated T cells 5 (NFAT5), and glycogen synthase kinase 3β (GSK3B), the risk haplotype had...
Figure 1: Association of the canonical interferon pathway with the IRF5 haplotype. Naive B cells were either exposed to EBV for 16 hours or left unstimulated. Whole-genome expression levels were compared between cells with the risk and protective haplotypes using the Illumina platform. Significant association of the data set with the canonical IFN pathway was discovered using Ingenuity Pathway Analysis (P < 0.01). Interactions between genes in the interferon pathway based on Ingenuity Pathway Analysis are shown for both the unstimulated (a) and the EBV-infected conditions (b). Blue gene symbols represent genes with relatively lower expression in the cells with the risk haplotype, while genes with red shading are upregulated in the risk cells. IFIT3: IFN-induced protein with tetratricopeptide repeats 3, SOCS: suppressor of cytokine signaling.

3.4. The Interferon Response Signature in Patients Depends on Haplotype. Lupus patients have a heightened interferon response signature in the peripheral blood [50–52]. This signature is heritable and is associated with the IRF5 risk haplotype [19, 26]. When we examined genes included in the interferon response signature, we found an interesting association with the IRF5 risk haplotype. Cells from the SLE patients with the risk haplotype had an interferon response signature under all conditions, whether exposed to EBV or not. However, the cells from SLE patients with the protective haplotype did not exhibit an interferon signature without EBV infection. The difference in expression of the interferon response genes between the unstimulated patient risk and the unstimulated patient protective cells was statistically significant (P = 0.011) (Figure 4). The risk haplotype cells derived from control individuals did not have heightened baseline expression of interferon response genes. After exposure to EBV, these cells developed an interferon response signature that was similar to that seen in the baseline and EBV-infected risk-haplotype lupus patients.

Interestingly, the patients with the protective haplotype did not develop a strong interferon response signature even after exposure to EBV, indicating that the IRF5 protective haplotype is dampening the response to interferon compared to the risk haplotype (Figure 4).

4. Discussion

The IRF5 gene has been associated with risk for lupus. These findings demonstrate that the lupus-associated polymorphisms in the IRF5 gene have wide-reaching effects on B-cell responses to infection. The gene sets that were enriched in the risk haplotypes included interferon-related sets, which is encouraging considering that the genotype being examined is IRF5. Multiple gene sets that are related to lupus were enriched in the cells with the risk haplotypes, including IFNα sets, interleukin- (IL-) 6, inflammation, proliferation, and monocyte and dendritic cell genes, in addition to the SLE-related interferon gene set. The finding that these gene sets are the most strongly enriched in the risk haplotype indicates that the IRF5 risk haplotype has a strong influence on interferon signaling and inflammation, processes that are at the core of SLE. The finding that the most enriched gene sets were associated with interferon and lupus also indicates
Table 3: Genes exhibiting differential expression between risk and nonrisk cells in the canonical pathways identified through Ingenuity pathways analysis. Fold up/down column is positive in the case that the gene expression is higher in the risk haplotype cells, and negative in the case that gene expression is higher in the nonrisk cells. PIK3CA: phosphoinositide-3-kinase, catalytic, alpha; RAC1: Ras-related C3 botulinum toxin substrate 1. *The IFNB1 gene is found in both the interferon and Toll-like receptor pathways.

<table>
<thead>
<tr>
<th>Gene pathway</th>
<th>Gene symbol</th>
<th>Mock infected 16 hours</th>
<th>EBV live virus infected 16 hours</th>
<th>Fold up/down</th>
<th>Fold up/down</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interferon</td>
<td>IFNB1*</td>
<td>104.95</td>
<td>3.89</td>
<td>0.04</td>
<td>−26.99</td>
</tr>
<tr>
<td></td>
<td>STAT1</td>
<td>71.08</td>
<td>52.84</td>
<td>0.74</td>
<td>−1.35</td>
</tr>
<tr>
<td></td>
<td>OAS1</td>
<td>56.04</td>
<td>84.55</td>
<td>1.51</td>
<td>1.51</td>
</tr>
<tr>
<td></td>
<td>MX1</td>
<td>3179.14</td>
<td>5161.02</td>
<td>1.62</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td>IFNAR2</td>
<td>959.14</td>
<td>1637.11</td>
<td>1.71</td>
<td>1.71</td>
</tr>
<tr>
<td></td>
<td>IFITM1</td>
<td>439.88</td>
<td>944.41</td>
<td>2.15</td>
<td>2.15</td>
</tr>
<tr>
<td>Toll-like receptor</td>
<td>IFNB1*</td>
<td>104.95</td>
<td>3.89</td>
<td>0.04</td>
<td>−26.99</td>
</tr>
<tr>
<td></td>
<td>FOS</td>
<td>116.48</td>
<td>53.84</td>
<td>0.46</td>
<td>−2.16</td>
</tr>
<tr>
<td></td>
<td>MYD88</td>
<td>281.53</td>
<td>152.57</td>
<td>0.54</td>
<td>−1.85</td>
</tr>
<tr>
<td></td>
<td>TNFAIP3</td>
<td>439.96</td>
<td>321.29</td>
<td>0.73</td>
<td>−1.37</td>
</tr>
<tr>
<td></td>
<td>TLR1</td>
<td>26.50</td>
<td>46.76</td>
<td>1.76</td>
<td>1.76</td>
</tr>
<tr>
<td></td>
<td>LY96 (MD-2)</td>
<td>560.79</td>
<td>1274.32</td>
<td>2.27</td>
<td>2.27</td>
</tr>
<tr>
<td></td>
<td>CD14</td>
<td>401.17</td>
<td>1272.53</td>
<td>3.17</td>
<td>3.17</td>
</tr>
<tr>
<td>B-Cell receptor</td>
<td>CD79B</td>
<td>2222.90</td>
<td>991.48</td>
<td>0.45</td>
<td>−2.24</td>
</tr>
<tr>
<td></td>
<td>CD79A</td>
<td>41.32</td>
<td>19.08</td>
<td>0.46</td>
<td>−2.17</td>
</tr>
<tr>
<td></td>
<td>RAC1</td>
<td>1407.29</td>
<td>930.84</td>
<td>0.66</td>
<td>−1.51</td>
</tr>
<tr>
<td></td>
<td>MAPK9</td>
<td>105.87</td>
<td>71.02</td>
<td>0.67</td>
<td>−1.49</td>
</tr>
<tr>
<td></td>
<td>AKT1</td>
<td>512.79</td>
<td>353.65</td>
<td>0.69</td>
<td>−1.45</td>
</tr>
<tr>
<td></td>
<td>NFkB2</td>
<td>6.44</td>
<td>4.85</td>
<td>0.75</td>
<td>−1.33</td>
</tr>
<tr>
<td></td>
<td>PIK3CA</td>
<td>25.86</td>
<td>40.05</td>
<td>1.55</td>
<td>1.55</td>
</tr>
<tr>
<td></td>
<td>Nfat5</td>
<td>123.74</td>
<td>222.63</td>
<td>1.80</td>
<td>1.80</td>
</tr>
</tbody>
</table>

that these results are unlikely to be false positives obtained by chance, since the variable being studied is an interferon-affecting gene.

The gene set enrichment analysis techniques that were used are valuable because they identify not only individual genes, but also how strongly pathways that include those genes and the interactions between them are affected by the experimental conditions. This allows a much broader look into gene networks than looking only at individual genes. These studies point to the IRF5 risk haplotype having a wide influence on interferon and inflammation. The results identify targets for future investigation into the function of the IRF5 polymorphisms as well as other genetic influences on lupus.

As was the case with the gene set enrichment analysis, the identification of the interferon and Toll-like receptor pathways through Illumina pathway analysis suggests that the results are robust, as these are pathways that would be expected to be modulated by the underlying IRF5 haplotype of the donor. Interferon alpha is an extremely important cytokine in lupus [53]. These studies suggest that the interferon alpha pathway is strongly affected by genetic variation in the IRF5 gene, and show multiple genes that could potentially be targets for understanding interferon in lupus or potential therapeutic targets. Toll-like receptors are involved in response to infection through the recognition of pathogen-associated molecular patterns. Additionally, Toll-like receptors are important in the pathogenesis of lupus. They are capable of recognizing endogenous nucleic acids in the context of immune complexes found in lupus patient sera, thereby stimulating dendritic cell maturation and interferon alpha production [25, 54–60], a process which also involves IRF5 itself [25]. IRF5 is a very interesting transcriptional regulator in that it acts as both an activator when homodimerized and blocks activation when heterodimerized with IRF7 [61, 62]. This mechanism of action may help to explain how some of these pathways can exhibit relative upregulation or downregulation depending on the other conditions in the cell.

One of the more unexpected findings of this study was the modulation of the B-cell receptor pathway by the IRF5 haplotype. The B-cell receptor is important in the recognition of antigen and the survival, maturation, and proliferation of B cells. B cells produce the autoantibodies
involved in lupus, as well as being important for antigen processing and presentation and T-cell activation. Differences in the activation threshold or other effects that may be seen with altered B-cell receptor gene expression may be very important to breaking self-tolerance or other aspects of B cell biology involved in SLE. Of particular interest in this respect is the recent finding that IRF5 controls antibody class switching to IgG2A, allowing lupus-like autoimmunity in mice [63]. The B-cell receptor and Toll-like receptor pathways are involved in antibody class switching, and the genes that were modulated by IRF5 variation in this study could represent mechanisms through which the IRF5 risk haplotype may contribute to class switching or other similar variations in humans.

The interferon response signature has been identified as a common feature in lupus. These studies examined how polymorphisms in the IRF5 gene affected the interferon response signature in both patients and controls. Interestingly, SLE patients with the risk haplotype demonstrated an interferon signature in both the infected and uninfected cells, while a strong interferon response was not found in the patients without the risk haplotype even when stimulated by EBV exposure. The controls with the risk haplotype lacked the interferon response signature in the basal state, but developed it after exposure to EBV, as would be expected. These findings suggest that the IRF5 risk haplotype is integral for the interferon response signature in both patients and controls. They also indicate that other factors contribute to a basal interferon response in lupus patients, since the IRF5 risk haplotype was not sufficient for the response signature to be present in the unstimulated control cells, as it was in the patients with the risk haplotype.

Cells were infected with EBV for two reasons. The first was to identify differences in gene expression patterns when
cells were stimulated with a biologically relevant trigger for interferon production. The second reason was to identify areas that may start to explain the differences in EBV infection and response in lupus. Gene expression was examined for genes in the three pathways found to be significant by ingenuity analysis. In several cases, (IFITM1, IFNAR2, LYP6, PIK3CA, NFAT5, and GSK3B) the baseline level of gene expression was higher in the risk cells, but after EBV infection, the gene expression was comparatively increased in the protective cells. In other genes, including CD79A, CD79B, STAT1, MyD88, and Fos, expression was lower in the risk cells but the difference diminished or reversed after EBV infection. Expression of one gene, TNFAIP3, was lower in the risk than in the protective haplotype subject unstimulated cells and was comparatively diminished further after EBV infection. These differences suggest several areas of investigation to understand differences in B cell biology in lupus and show that the IRF5 haplotype affects multiple genes related to EBV infection and response.

Although a detailed analysis of each gene involved in these pathways is beyond the scope of this paper, the genes with expression differences between the risk and protective haplotypes are suggestive in several instances. One of the genes identified with promise to affect lupus is TNFAIP3. This gene is a transcription factor that is produced in lupus and other autoimmune and inflammatory diseases, and it is often suppressed in tumors, especially lymphomas [66, 72, 73]. Other promising genes identified by these experiments include STAT4, IFITM1, and IFNAR2,
Figure 4: Dependence of the interferon response signature on IRF5 haplotype in SLE patients. Interferon signature genes identified in gene set enrichment analysis were compared between patients and controls, and between lupus patients with the IRF5 risk and nonrisk haplotypes. Each column represents expression of genes of interest in individual subjects. (a) Interferon signature genes of 3 IRF5 nonrisk and 2 IRF5 risk SLE patients. Gene expression comparisons are between IRF5 risk and nonrisk SLE patients, either nonstimulated or EBV infected. (b) Gene set enrichment of 3 control IRF5 risk haplotype individuals and 3 SLE patient IRF5 risk haplotype individuals. Gene expression comparisons between IRF5 risk patients and controls from nonstimulated and EBV-infected B cells. ADAR: adenosine deaminase, BAK: Bcl2-antagonist/killer, BTG: B cell translocation gene, C1S: complement component 1S, CASP: caspase, CEBPD: CCAAT/enhancer-binding protein δ, eIF2B: eukaryotic translation initiation factor 2B, FOSL: Fos-related antigen, HADHB: hydroxyacyl-Coenzyme A dehydrogenase/3-ketoacyl-Coenzyme A thiolase/enoyl-Coenzyme A hydratase β subunit, HLA: human leukocyte antigen, PHLDA: pleckstrin homology-like domain family A, PMAIP1: phorbol-12-myristate-13-acetate-induced protein 1 (Noxa), PRAME: melanoma antigen preferentially expressed in tumors, RBBP: retinoblastoma binding protein, SF3A: splicing factor 3A, TRIM: tripartite motif, VAT: vesicle amine transport, and XRCC: X-ray repair cross-complementing (Ku70).

which are all involved in the response to interferon, and several B cell signaling genes, including NFAT5, GSK3B, and NFκB2.

Although EBV was used in part to simulate an infected state in B cells, EBV itself could be involved in the etiology of lupus by affecting several pathways. The three pathways identified here are all involved in EBV infection. EBV may stimulate these pathways through several mechanisms, including both infection and binding of virions to the receptors involved in these pathways. Although the effect of EBV infection on differential gene expression was somewhat variable, for many of the genes examined in this study there was overexpression in the risk cells, which subsequently diminished after EBV infection. This pattern, as well as that seen with the interferon response signature, suggests that the IRF5 risk haplotype makes these cells appear more activated in the resting state. Because of this heightened activation state, there is less difference in the response to EBV infection in the risk cells, with the nonrisk cells often catching up to or passing the risk cells in expression of several genes following viral infection. An activated basal state would be likely to promote inappropriate cellular responses and possibly heightened sensitivity to self-antigens, including those recognized by TLRs.

These findings identify several key pathways that are affected by the IRF5 risk haplotype and are involved in the B cell response to antigen stimulation and viral infection. Many of the genes involved in these pathways have definite potential to alter the response to EBV infection and affect the development of lupus. These merit further investigation. Since all of these pathways are likely to be involved in the development of lupus, further comparison of these pathways in other cell types such as plasmacytoid dendritic cells will be beneficial to understanding the origins and pathogenesis of lupus. It will also be beneficial to examine more closely the role of EBV in regulating expression of these genes, through the use of EBV mutants, and to dissect the role of IRF5 in each pathway and gene set identified.
Acknowledgments

This work was funded by the National Institutes of Health Grants P30 AR053483 (J. James and J. Guthridge), P20 RR015577 (J. James), N01 AI50026 (J. James and J. Guthridge), R01 AI063274 and R01 AR056360 (P. Gaffney), R3724717 (J. Harley), AR042460 (J. Harley), P01 AI083194 (J. Harley), P20 RR020143 (P. Gaffney), R01 DE018209 (J. Harley), and the Office for Human Research Studies award no. HR08–037 from the Oklahoma Center for the Advancement of Science & Technology (J. Guthridge).

References


5. References


51. Sharma RP, He Q, Riley RT. Lupus-prone NZBWFl/J mice, defective in cytokine signaling, are resistant to fumonisin hepatotoxicity despite accumulation of liver sphinganine. Toxicology, 2005; 216:59-71.


94. Lofgren SE, Yin H, Delgado-Vega AM, Sanchez E, Lewen S, Pons-Estel BA et al. Promoter insertion/deletion in the IRF5 gene is highly associated with susceptibility to systemic lupus
erythematous in distinct populations, but exerts a modest effect on gene expression in peripheral blood mononuclear cells. J Rheumatol, 2010; 37:574-8.


223. Bave U, Magnusson M, Eloranta ML, Perers A, Alm GV, Ronnblom L. FcγRIIa is expressed on natural IFNα -producing cells (plasmacytoid dendritic cells) and is required for the IFNα production induced by apoptotic cells combined with lupus IgG. J Immunol, 2003; 171:3296-302.


239. Grievink H, Stowell K. Allele-specific differences in ryanodine receptor 1 mRNA expression levels may contribute to phenotypic variability in malignant hyperthermia. Orphanet J Rare Dis, 2010; 5:10.


