The Genetic Basis of Systemic Lupus Erythematosus

Bohan Zheng

Abstract:

Introduction: Autoimmune diseases, such as Systemic Lupus Erythematosus (hereafter SLE), often are linked to many variable factors. Previously, it had been impossible to determine whether genetic susceptibility influenced such diseases, but that has become possible with advancements in GWAS and databases. In this research, we analyzed the genetic inheritance of SLE, and whether certain genes may constitute a higher chance of developing the disease.

Methods: Through multiple gene-wide association studies, 36 plausible SNPs were found. Then, a series of analyses including gene linkage maps, gene expression, and gene function research were conducted with the help of database tools. Results were concluded and presented through an assortment of Tables and Figures.

Results: All 36 SNPs had an OR value higher than 1.2, proving that they did impact susceptibility to SLE. Results showed linkage between the *BLK* and *BANK1* genes in susceptibility to SLE. A majority of SNPs which caused SLE were located in intronic areas, suggesting that they presented transcription factors that controlled a few exonic SNPs.

Conclusion: This study showed the assortment of SNPs that could affect SLE susceptibility, connection between gene function, as well as the linkage between the genes involved. Further experiments will provide information regarding the mechanisms by which genes regulate cytokines and symptoms of SLE.

Key Words:

Systemic Lupus Erythematosus, SNP, Single Nucleotide Polymorphism, BLK, BANK1, STAT4, HLA, MHC

1. Introduction to SLE

Systemic Lupus Erythematosus (SLE) is an autoimmune disease in which the immune system attacks its own tissues, causing inflammation and tissue damage in affected organs. These may be the skin, joints, brain, lungs, kidneys, and blood vessels. The disease involved antibodies assuming the body's own functional cells as targets.

SLE is often mistaken for other illnesses. Symptoms of SLE are similar to that of many other diseases, such as fevers, malaise, muscle pain, and fatigue. These symptoms are not part of the diagnostic process of SLE, however when put in combination with one another, they may be a sign of SLE. Most common symptoms include thick, scaly skin rashes, joint pain and anemia. SLE is also found in women more than men at a disproportionate ratio. Global rates of SLE are also comparatively high.

SLE runs in families, and studies do show that the rate of heredity of SLE is over 66%. However, multiple genes may influence a person's chance to develop lupus. As will be elaborated upon further on, these genes include, but are not limited to HLA Class I and II, STAT4, BANK1, and *BLK*. A number of studies have been conducted in the past to research the causes of SLE, a few of which focus on the genetic inheritance of the disease. Though these studies combine to provide a variety of causes, it has been concluded that susceptibility to SLE is not caused by a mutation in a certain gene. Rather, SNPs found in multiples genes are all related to SLE, and they collectively influence the rate at which SLE occurs.

This study performed a systemic analysis of SNPs and genes associated with SLE, using primarily the GWAS system. Over thirty SNPs, and a similar number of genes were involved in this study, and by analyzing their effects on human physiology and disease occurrence, this study sims to produce a systematic report on genetics causes of SLE. This process heavily involved looking at linkage between genes and how they worked together to influence susceptibility to SLE.

2. Materials and Methods

Multiple resources and databases were utilized in order to investigate the genetic inheritance of SLE. In order to gather a list of SNPs to study, GWAS Central (https://www.gwascentral.org/), a Gene-Wide Association Study database was used. GWAS Central provided a compiled summary of findings from nearly all gene association studies, and was a useful tool in finding SNPs related to SLE. $-\log p \geq 2$ (p-value <= 0.01) was used as criteria, such that the results were statistically significant and had positive correlation with SLE, and not due to chance. Results from seven studies appeared, though one was rejected due to it having over 400 SNPs, all with comparatively low OR index. An OR value is an index that depicts the odds ratio between two values, and a higher OR value corresponds to a likelier correlation between this SNP and susceptibility to SLE.

After above results were collected, a table was organized to summarize the details regarding these SNPs. In this process, wANNOVER (https://wannovar.wglab.org/) was utilized to functionally annotate all SNPs collected in this study. wANNOVER provided an assortment of information including the altered nucleotide, possibility of its mutation, as well as impact on susceptibility to disease.

Definition and annotation of genes themselves were completed through two resources, GeneCards (https://www.genecards.org/) and NIH's GenBank (www.ncbi.nlm.nih.gov/genbank/). The former provided basic information such as the full name, location, and function of the gene, while the latter gave details in regards to specific nucleotides by using BLAST. Then, these genes were filtered down manually such that localized, or undefined sequences were deleted. Furthermore, DAVID (https://david.ncifcrf.gov/tools.jsp), an analysis wizard that provided bioinformatics information, as well as correlations between genes, pathways, structures, and bodily functions.

GTEx (https://gtexportal.org/home/) then provided information regarding gene expression, expression location, and regulation. This portal was used to create charts in regards to the median expression of products of certain genes within the human body.

In this process, the UCSC Genome Browser (https://genome.ucsc.edu/) helped interactively visualize genome data, and the specific location of SNPs in their genes, SNPs from other GWAS studies, and portions of genes helpful to the study being conducted.

Last but not least, STRING (https://string-db.org/), a database used to create functional protein networks aided the process of assessing relations between certain proteins, as well as all genes involved from a macroscopic view.

3. Results

The chart below is based on results from GWAS. A total of 42 SNPs were found in relation to a higher susceptibility to SLE. All 42 SNPs have an OR of 1.2 or higher, meaning that they should a positive correlation to an enhanced risk of SLE.

A simple, early analysis of this data results in a majority of intergenic and intronic function, while 4 are exonic. To be exonic means that the genes are not spliced out during RNA processing, and form a certain protein, and to be intronic means that these genes may play a role in the regulation of other genes. From a chromosomal perspective, many of these SNPs seem to be located on either chromosome 2's STAT4 gene, chromosome 4's BANK1 gene, and chromosome 6's assortment of HLA genes. Further discussion will occur in regards to the functions of these genes.

ACCESSION	CHR	START	REF	ALT	FUNCTION	GENE	OR
RS6695567	1	53629585	G	A/C/T	intergenic	SLC1A7	1.32
RS525410	1	183176430	A	G	intronic	LAMC2	1.33
RS10737562	1	189723623	A	G	intergenic	BRINP3	1.3
RS17039212	2	49955497	C	A / G	intergenic	FSHR	1.72
RS918959	2	181513729	G	A	intergenic	CWC22	1.23
RS3821236	2	191902758	G	A	intronic	STAT4	1.49
RS7574865	2	191964633	T	G	intronic	STAT4	1.77
RS16841441	3	99340297	G	A	ncRNA_intronic	MIR548G	1.64
RS13138252	4	23757800	G	A / T	intergenic	GBA3	1.24
RS4522865	4	102715888	G	A / T	intronic	BANK1	1.33
RS4572885	4	102735513	T	A	intronic	BANK1	1.45
RS17266594	4	102750922	T	C	intronic	BANK1	1.56
RS10516487	4	102751076	G	A / T	exonic	BANK1	1.42

RS10516486	4	102751276	C	T	exonic	BANK1	1.75
RS17200824	4	102752589	A	G	intronic	BANK1	1.45
RS10516482	4	102780170	C	A	intronic	BANK1	1.33
RS10516483	4	102791905	C	G	intronic	BANK1	1.43
RS3733197	4	102839287	G	A	exonic	BANK1	1.56
RS4956211	4	109723126	G	A	intergenic	ETNPPL	1.71
RS2431697	5	159879978	T	C	intergenic	PTTG1	1.53
RS1150754	6	32050758	C	A / T	intronic	TNXB	1.44
RS2187668	6	32605884	C	T	intronic	HLA-DQA1, HLA-A	1.62
RS2647012	6	32664458	T	C	intergenic	HLA-DQB1	1.43
RS2301271	6	32725193	A	G	intronic	HLA-DQB2, HLA-B	1.71
RS5029939	6	138195723	C	G	intronic	TNFAIP3	1.54
RS1419842	7	34355100	C	A / T	intergenic	BMPER	1.44
RS10488631	7	128594183	T	C	downstream	TNPO3	1.74
RS12531711	7	128617466	A	C / G	intronic	TNPO3	1.44
RS2736340	8	11343973	C	T	intergenic	FAM167A;BLK	1.64
RS2618476	8	11352541	T	C	intronic	BLK	1.73
RS11101442	10	49936336	C	A / T	intronic	WDFY4	1.64
RS4963128	11	589564	T	C	intronic	PHRF1	1.21
RS7927370	11	55136219	C	T	exonic	OR4A15	1.42
RS7329174	13	41558110	A	G	intronic	ELF1	1.28
RS9888739	16	31313253	C	T	intronic	ITGAM	1.8
RS11150610	16	31334236	C	A	intronic	ITGAM	1.71
RS12949531	17	13733806	C	G / T	intergenic	HS3ST3A1	1.33
RS6049839	20	2518565	G	T	intronic	TMC2	1.37
RS6086678	20	8853275	A	T	intronic	PLCB1	1.36
RS5754217	22	21939675	G	T	intronic	UBE2L3	1.38

Table 1. SNPs and Genes related to SLE, according to GWAS

After this, DAVID was utilized in order to analyze the processes and pathways related to the aforementioned genes. The most common results, with the smallest p-values mostly involved the immune system and the MHC (major histocompatibility complex). SLE is an immune disease, which further confirms these results. Specifically, SLE is caused by an assortment of hyperactive B cells, which are more prone to polyclonal activation, and T cells increased in peripheral blood, as well as phagocytic cells that are unable to bind or process complex processes efficiently. Cytokines that cause the disease also seem to be involved by certain genes that SNPs point that, including BLK's IL-12.

CATEGORY	TERM	PERCENTAGE	P-VALUE
COMPONENT	MHC II	18.2%	1.1E-10
TERM	Antigen Processing	18.2%	7.9E-9
TERM	IGc1	18.2%	2.3E-7
PATHWAY	Graft-vs-host	15.2%	2.8E-6
PATHWAY	Type I diabetes	15.2%	3.1E-6
PATHWAY	Autoimmune thyroid disease	15.2%	7.1E-6
COMPONENT	Lysosome	21.2%	2.3E-5
PATHWAY	Phagosome	18.2%	2.6E-5
PATHWAY	Cell Adhesion Molecules	18.2%	3.0E-5
COMPONENT	Endosome	21.2%	4.2E-4
PATHWAY	Immune Response	18.2%	7.6E-4
DISEASE	Susceptibility to SLE	9.1%	1.2E-3
PROCESS	Immunity	21.2%	1.0E-2
TERM	Regulation of B Cell Activation	9.1%	2.2E-2
DISEASE	SLE	9.1%	4.3E-2
COMPONENT	Lysosome	6.9%	9.2E-2
FUNCTION	Glycosidase	10.3%	9.4E-2

Table 2. Processes and pathways shared by the genes mentioned

After this, GeneCards and NIH's GenBank were utilized in order to research the function of each specific gene, thus presenting their correlation to and pathologies with SLE. 33 genes in total were involved in this analysis, 6 of which were part of the HLA (human leukocyte antigen) system. Genes such as BANK1, STAT4, and BLK are presumed to increase susceptibility to SLE. Deeper analysis of this data will be presented after the chart.

Gene Name	Full Name	Location	Function
SLC1A7	Solute Carrier Family 1 Member 7	Plasma Membrane	This gene enables L-glutamate transmembrane transporter activity.
LAMC2	Laminin Subunit Gamma 2	Secreted	This gene laminin mediates the attachment, migration and organization of cells into tissues during embryonic development.
BRINP3	BMP/Retinoic Acid Inducible Neural Specific 3	Endoplasmic Reticulum	This gene inhibits neuronal cell proliferation by negative regulation of the cell cycle transition, and may cause a heightened risk of pituitary tumor.
FSHR	Follicle Stimulating Hormone Receptor	Cell Membrane	This gene activates a G protein-coupled receptor for the follicle-stimulating hormone.
CWC22	CWC22 Spliceosome Associated Protein Homolog	Nucleus	This gene enables RNA binding activity and is involved in mRNA splicing. Diseases associated with this gene include Hard Palate Cancer.
STAT4	Signal Transducer and Activator of Transcription 4	Chromatin (Nucleus)	This gene encodes a member of the STAT family of transcription factors. This protein mediates responses to IL-12 in lymphocytes, and regulates the differentiation of T helper cells. Mutations may be associated with SLE.
NEU2	Neuraminidase 2	Cytoplasm	This gene belongs to a family of glycohydrolytic enzymes which remove sialic acid residues from glycoproteins and glycolipids. Diseases include galactosialidosis and glycoproteinosis.
MIR548G	MicroRNA 548g	Nucleus	This is a non-coding RNA.
GBA3	Glucosyl ceramidase beta 3	Cytoplasm	The protein encoded by this gene is a cytosolic enzyme that can hydrolyze several types of glycosides. This gene is a polymorphic pseudogene.
BANK1	B cell Scaffold Protein with Ankyrin Repeats 1	Cytoplasm	This gene encodes a B-cell-specific scaffold protein that functions in B-cell receptor-induced calcium mobilization from intracellular stores. Polymorphisms may be associated with SLE.
ETNPPL	Ethanolamine-Phosphate Phospho-lyase	Mitochondria	This gene enables Ethanolamine-Phosphate Phospho-lyase activity, breaking down phosphoethanolamine.
PTTG1	PTTG1 Regulator of Sister Chromatid Separation	Nucleus	This gene encodes a regulatory protein, which plays a central role in chromosome stability, in the p53/TP53 pathway, and DNA repair. Diseases associated include pituitary tumors.
TNXB	Tenascin XB	Secreted	This gene encodes a protein that mediates interactions between cells and the extracellular matrix. Mutations may lead to epithelial tumors.
HLA-DQA1	Major Histocompatibility Complex, Class II, DQ alpha 1	Endoplasmic Reticulum	This gene is a heterodimer consisting of an alpha (DQA) and a beta chain (DQB), both anchored in the membrane. It plays a central role in the immune system by presenting peptides derived from extracellular proteins.
HLA-DQA2	Major histocompatibility complex, class II, DQ alpha 2	Cell Membrane	HLA-DQA2 plays a central role in the immune system by presenting peptides derived from extracellular proteins.

HLA-A	Major Histocompatibility Complex, Class I, A	Cell Membrane	HLA-A plays a central role in the immune system by presenting peptides derived from the endoplasmic reticulum lumen so that they can be recognized by cytotoxic T cells.
HLA-B	Major Histocompatibility Complex, Class I, B	Cell Membrane	HLA-B plays a central role in the immune system by presenting peptides derived from the endoplasmic reticulum lumen
HLA-DQB1	Major Histocompatibility Complex, Class II, DQ beta 1	Cell Membrane	HLA- DQB1 plays a central role in the immune system by presenting peptides derived from extracellular proteins.
HLA-DQB2	Major Histocompatibility Complex, Class II, DQ beta 2	Endoplasmic Reticulum	This gene is a heterodimer consisting of an alpha (DQA) and a beta chain (DQB), both anchored in the membrane. It plays a central role in the immune system by presenting peptides derived from extracellular proteins.
TNFAIP3	TNF alpha induced protein 3	Lysosome	This gene's expression is rapidly induced by the tumor necrosis factor. The encoded protein, which has both ubiquitin ligase and deubiquitinase activities, is involved in the cytokine-mediated immune and inflammatory responses.
BMPER	BMP Binding Endothelial Regulator	Secreted	This gene encodes a secreted protein that interacts with, and inhibits bone morphogenetic protein (BMP) function.
TNPO3	Transportin 3	Nucleus	The protein encoded by this gene is a nuclear import receptor for serine/arginine-rich (SR) proteins such as the splicing factors SFRS1 and SFRS2. The encoded protein has also been shown to be involved in HIV-1 infection.
FAM167A	Family With Sequence Similarity 167 Member A	-	Diseases associated with this gene include Diabetes and Erythema.
BLK	BLK Proto-Oncogene, Src Family Tyrosine Kinase	Cell Membrane	This gene encodes a nonreceptor tyrosine-kinase of the src family of proto-oncogenes that are typically involved in cell proliferation and differentiation. The protein has a role in B-cell receptor signaling and B-cell development.
WDFY4	WDFY Family Member 4	Endoplasmic Reticulum	This gene is predicted to act upstream of or within with a positive effect on CD8-positive, alpha-beta T cell activation.
PHRF1	PHD And Ring Finger Domains 1	Nuclear Membrane	This gene is predicted to be involved in mRNA processing and transcription by RNA polymerase II. This gene is associated with SLE.
OR4A15	Olfactory Receptor Family 4 Subfamily A Member 15	Cell Membrane	Olfactory receptors interact with odorant molecules in the nose, to initiate a neuronal response that triggers the perception of a smell. The olfactory receptor proteins are members of a large family of G-protein-coupled receptors arising from single coding-exon genes.
ELF1	E74 like ETS Transcription Factor 1	Chromatin (Nucleus)	This gene encodes an E26 transformation-specific related transcription factor. The encoded protein is primarily expressed in lymphoid cells and acts as both an

			enhancer and a repressor to regulate transcription of various genes.
ITGAM	Integrin Subunit Alpha M	Cell Membrane	This gene encodes the integrin alpha M chain. The alpha M beta 2 integrin is important in the adherence of neutrophils and monocytes to stimulated endothelium, and also in the phagocytosis of complement coated particles.
HS3ST3A1	Heparan Sulfate-Glucosamine 3-Sulfotransferase 3A1	Golgi Apparatus	Heparan sulfate biosynthetic enzymes are key components in generating a myriad of distinct heparan sulfate fine structures that carry out multiple biologic activities. This gene is widely expressed, with the most abundant expression in liver and placenta.
TMC2	Transmembrane Channel like 2	Cell Membrane	This gene encodes a transmembrane protein that is necessary for mechanotransduction in cochlear hair cells of the inner ear. Mutations in this gene may underlie hereditary disorders of balance and hearing.
PLCB1	Phospholipase C beta 1	Cell Membrane	The protein encoded catalyzes a reaction that plays an important role in the intracellular transduction of many extracellular signals.
UBE2L3	Ubiquitin-Conjugating Enzyme E2 L3	Nucleus	This gene encodes for Ubiquitin-Conjugating Enzyme E2, and is involved in the selective degradation of short-lived and abnormal proteins.

Table 3. Genes related to SNP, and their functions.

In regards to the 33 genes presented above; multiple observations were made. Firstly, multiple genes were part of the HLA gene system, which certifies that SLE is an immune system disease. These genes include *HLA-DQA1, HLA-DQA2, HLA-A, HLA-B, HLA-DQB1, HLA-DQB2,* which in combination, play a central role in the immune system by expressing peptides from both intracellular and extracellular spaces. As per multiple studies, variations and alterations in the MHC locus are associated with a wide range of autoimmune diseases, which includes SLE in this spectrum. In recent years, studies have been conducted to show the link of MHC Class II proteins with patients of specific populations, but a link to explain how this occurs has not been found. The role that HLA genes play in autoantibody expression has been researched, and indicates the activation of autoaggressive B cells and breakdown of tolerance to self-antigens, a common pathology of SLE

Specific Genes and their Impacts

Genes worth noting, especially after viewing their respective functions include *BANK1, PHRF1, BLK, STAT4, ITGAM* and *ELF1.* Detailed elaborations, predictions and information are as listed below:

1 ***BLK***: Studies have shown that SNPs in BLK impair the phosphorylation of BANK1, signifying the connection between the aforementioned genes. BLK mutation may cause an impairment in kinase function, as well as other kinase inabilities. In addition, BLK kinase activity caused by SNPs impaired interferon-beta repression, which in turn led to the symptoms of SLE.

2 ***BANK1***: The BANK1 gene is predicted to have correlations with SLE susceptibility. The BLK gene and BANK1 genes are tightly connected, which will be observed in the gene map later presented. This study predicts that variants of the BANK1 gene, which may be caused by SNPs of the BLK gene cause altered B cell signaling and an enhanced level of memory B cell development. According to studies previously performed, SNPs in BANK1 enhance nuclear IRF5 localizations, which ultimately leads to increased T1 IFN activity in B cells, a common feature in the majority of SLE patients.

3 ***STAT4***: STAT4 is involved in interleukin-12 and T helper cell expression and production. Both of these are in tight correlation with SLE, especially as patients with SLE showed significantly higher levels of IL-12, a cytokine that inhibits IL-10. The misbalance between the aforementioned cytokines plays an important role in cellular immune responses in SLE patients. Likewise, this leads to increased T1 IFN activity in B cells.

4 ***ITGAM***: Previously, not many studies had linked this gene to SLE, but this study observed the role *ITGAM* played in phagocytosis. SLE is often connected with abnormal phagocytic function, where phagocytic cells cannot bind or process immune complexes efficiently, as well as high adhesion molecule expression in endothelial cells. Both these functions are mediated by *ITGAM*, thus the correlation between gene and disease is observed and will be further explored.

5 ***PHRF1***: This gene is connected to the production and expression of interferon-alpha, through controlling the transcription factor IRF7, which is similar to the previously

mentioned IRF5. PHRF1 is located 23kb telomeric towards this transcription factor, and causes high levels of interferon-alpha to be synthesized.

6 ***ELF1***: This gene is expressed in lymphoid cells and may play a role in both IL-2 cytokine synthesis and T-cell receptor signaling. SLE is often connected with decreased IL-2 production in activated peripheral blood T cells. *ELF1's* functions may provide a link to why its SNP could cause a higher susceptibility to SLE.

A Link exists between the BLK and BANK1 genes

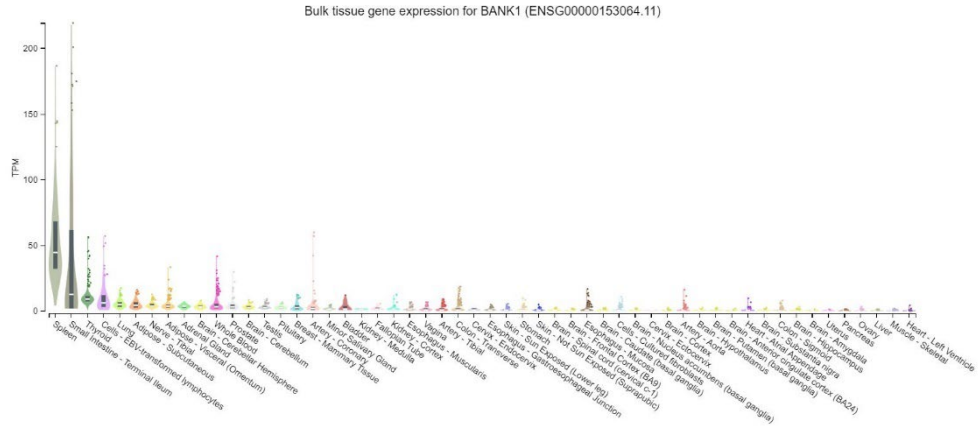

Figure 4: Bulk tissue expression for BLK Gene

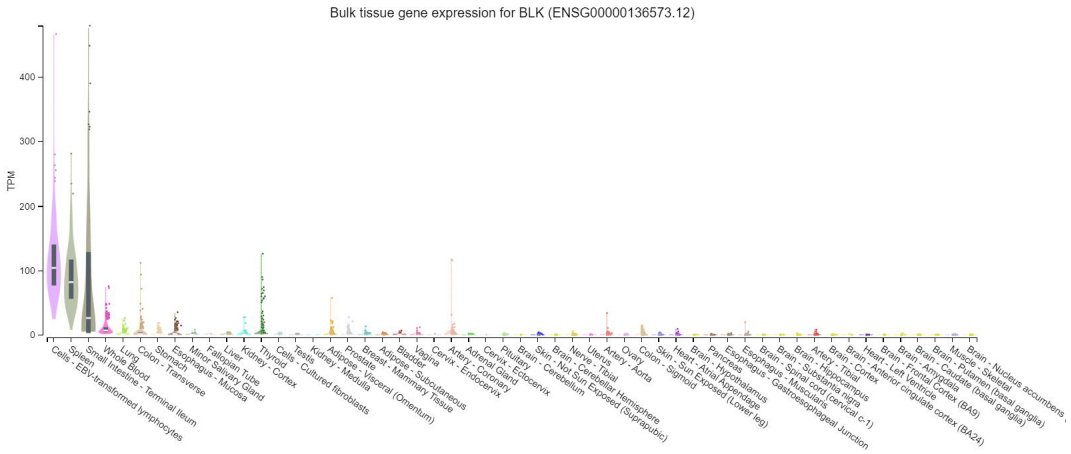

Figure 5: Bulk tissue expression for BANK1 Gene

As shown above in the two Figures, the BLK and BANK1 gene are commonly expressed in the same tissues, including the spleen, small intestine – terminal ileum, and EBV-transformed lymphocytes. Thus, it is hypothesized that BLK is somehow

related to BANK1 and its expression, since the transcription factor encoded by the BANK1 gene contributes to IRF5 expression, which then causes B cell proliferation, and common SLE symptoms.

The methods by which the BLK gene influences susceptibility to SLE include:

1. An impaired ability of phosphorylating BANK1. As BLK and BANK1 are linked, BLK genes with SNPs are thus unable to regulate the expression of BANK1.
2. The mutated BLK gene causes an inability in expression of IRF5 and type-I IFN in human B cell lines. Activation of IRF5 is a leading cause of SLE, and B lymphocytes' secretion of cytokines is often caused by an enhanced amount of IRF5. IRF5 is needed by B cells in order to secrete IgG2a, an autoantibody that acts against nuclear DNA functions, an important feature of SLE. An upregulation of type-I interferon, caused by the BLK mutation, is also a key feature of SLE response to autoantigens.
3. The mutated BANK1 gene, when simultaneously working in epistasis with the mutated BLK gene, is impaired in its regulation process. However, BANK1 including SNPs fails to repress IRF5 nearly as well as original BANK1 genes, thus activating IRF5 and leading to SLE symptoms.
4. As presented in Figure 4 and 5, both BLK and BANK1 are highly expressed in the spleen, blood, lymphocytes and other lymphoid tissue locations. In these locations, these genes tend to be located in plasmacytoid dendritic cells. Both genes have been proven to lead to an inability to suppress T1 IFN genes. In these cells specifically, which function to produce cytokines, an impaired suppression also occurs, which leads to susceptibility to SLE.

Furthermore, the UCSC genome browser was utilized in order to analyze the specific locations of certain SNPs, and how these influenced the expression and connection between genes. Once these results were collected, another observation was made, which will be fully elaborated on in the next subsection.

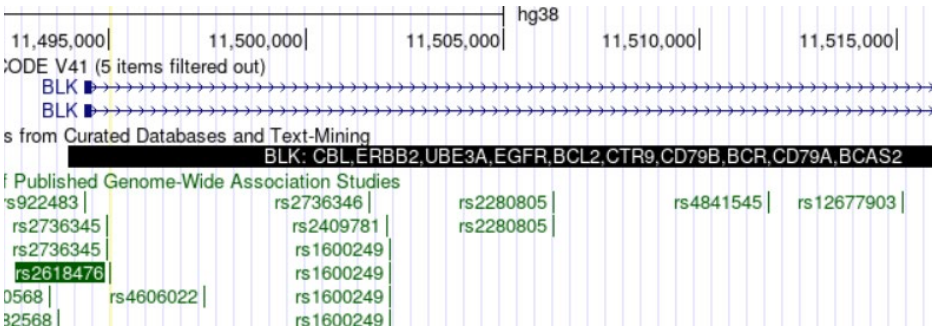

Figure 6: rs2618476 of the BLK gene is intronic

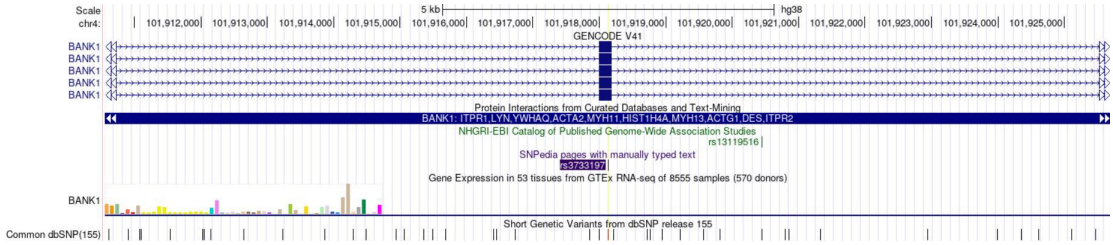

Figure 7: rs3733197 of the BANK1 gene is exonic

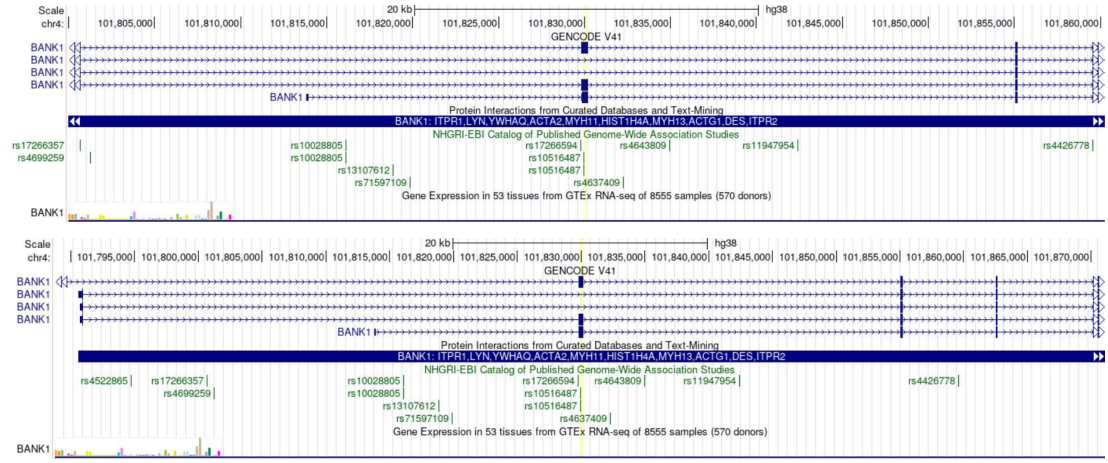

Figure 8 & 9: rs10516486 and rs10516487 of the BANK1 gene are exonic

Predominantly intron-based location of SNPs

After viewing the data for specific BLK and BANK1 genes, we were able to conclude that few of genes from the GWAS results were exonic. It was further predicted that much of these intronic genes could enhance, or regulate the expression of exonic genes, or other genes of the immune system not involved in the GWAS. Below are examples of intronic, intergenic and downstream SNPs.

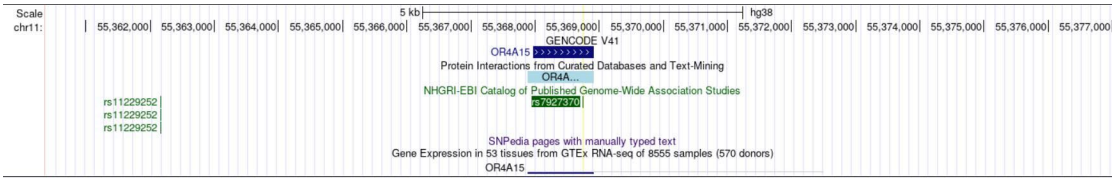

Figure 11: rs7927370 of the OR4A15 gene is the only other exonic SNP

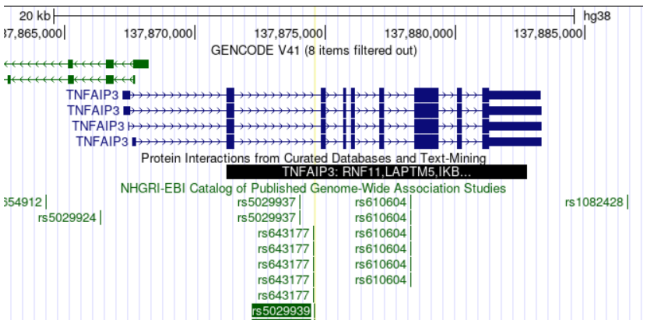

Figure 12: rs5029939 of the TNFAIP3 gene is intronic

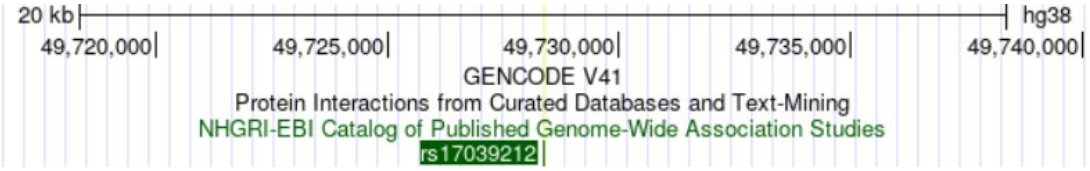

Figure 13: rs17039212 is intergenic.

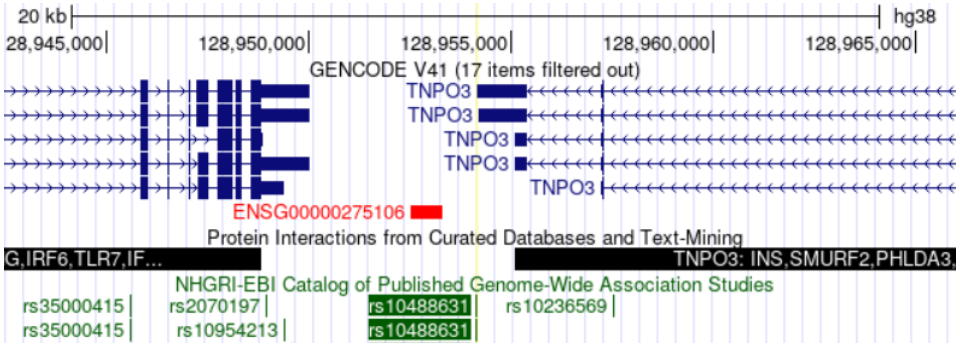

Figure 14: rs10488631 is downstream.

Gene Linkage Analysis

STRING was then utilized as a tool to form a detailed gene map considering the genes involved. It was observed that some genes had significantly more connections than others, and that ITGAM functioned as a connector between multiple gene groups. For example, the HLA (MHC II Class genes) had close interactions with one another, forming the collection of genes at the right top corner of the map.

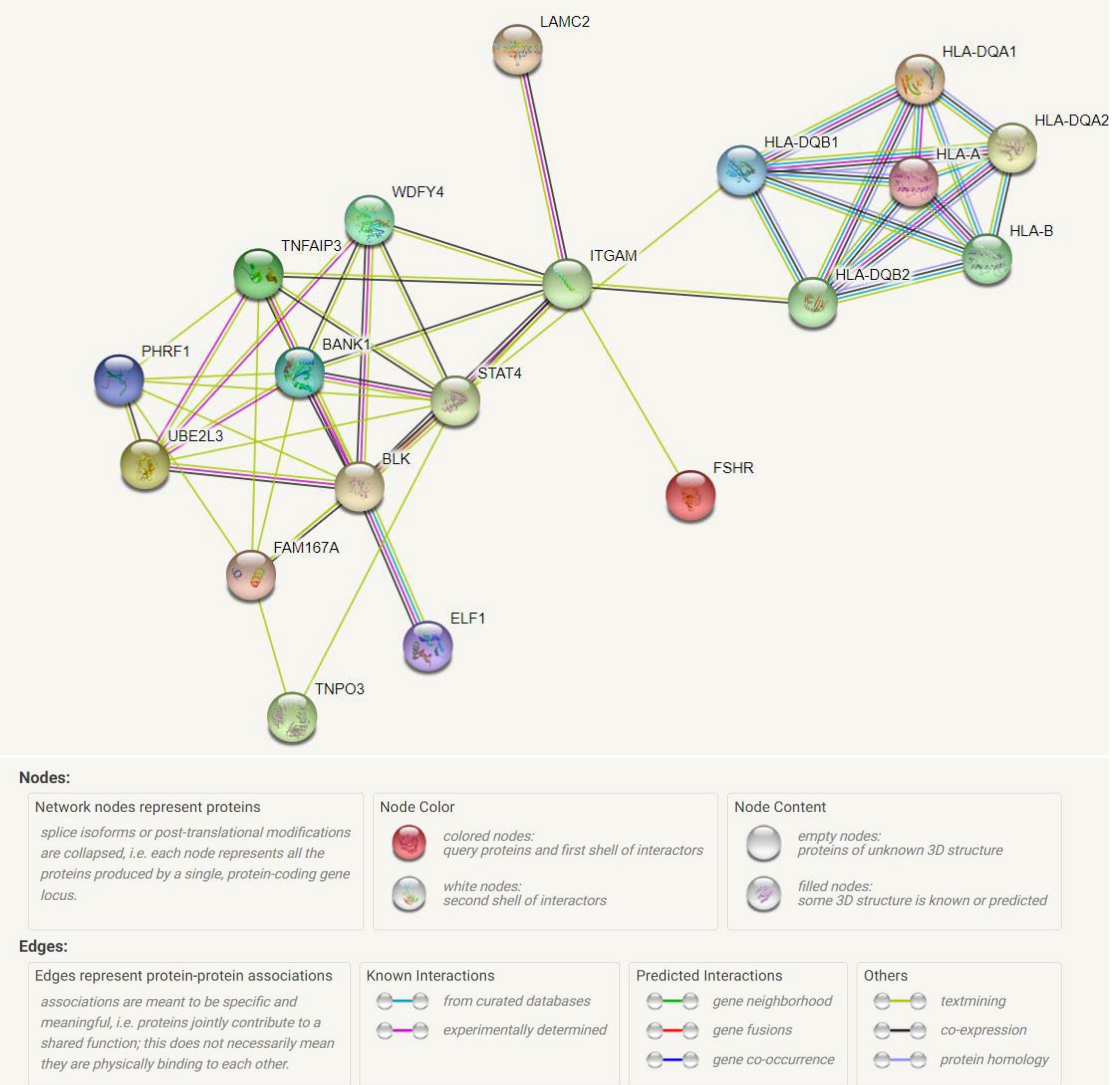

Figure 15: Simplified Gene Map of the Genes involved

As aforementioned, BLK and BANK1 had close connections as they were able to interact and co-express. However, STAT4, which is also a gene in correlation to SLE, also co-expressed with BLK and BANK1, which means that they control the expression of one another through transcription factors. A rise in production of the BLK

transcription factor, for example, causes an inhibition of BANK1 products.

Another noticeable factor was ITGAM's positioning in bridging the MHC Class II genes and other genes, which aren't comparatively as connected. After significant research, it was found that ITGAM had strong control over BLK transcription factor synthesis and kinase activity. ITGAM was also involved in pathways such as the dendritic cell pathway and phagocytosis pathway, both key to SLE development and functioning.

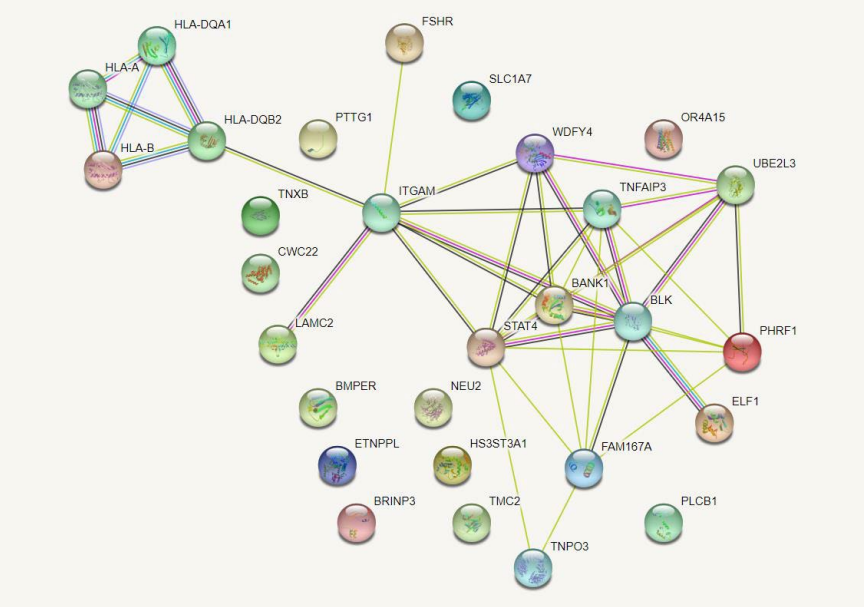

Figure 16: A non-simplified version of Figure 2.15

MHC Class II genes, and their role?

MHC genes, including HLA-A, HLA-B, HLA-DQA1, HLA-DQA2, HLA-DQB1, HLA-DQB2. HLA stands for Human Leukocyte Antigens, and play a central part in maintaining stability of the immune system. In patients with mutated genes, C2 and C4 proteins are under-expressed. Simultaneously, an antinuclear autoantibody called Ro/SSA, present in around 50% of SLE cases, occurred in conjunction with these mutated genes. It is known that MHC genes also monitor the production and expression of cytokines, including IL-2, IL-6, and IL-12, but the mechanism by which this occurs is unknown and requires further laboratory investigation.

4. Discussion

Since autoimmune diseases may have a great number of susceptibility factors and causes, genetics causes may take up only a small portion of how a diseases' phenotypes are developed. Thus, this study may not be a fully detailed description as to how, and which genetic mutations fully influence SLE.

The link between BLK and BANK1 genes were previously known, but deepened in understanding through the process of this study. The mechanism through which BLK and BANK1 are correlated and co-expressed, as well as interleukin activation were discovered. However, through more intense and experimental laboratory experiments can their specific transcription factor expression mechanisms be understood.

Links between certain genes that influence multiple genetic diseases were also previously unknown, such as that of ITGAM and BLK, and BLK and STAT4, which are all considered genes that increase susceptibility to SLE. These genes encode transcription factors which influence the translation of the products of other genes.

This study utilized primarily online resources to conduct. If laboratory resources were involved, a clearer understanding of processes and mechanisms would occur and be included.

5. Conclusion

This study investigated links between genetic mutations and Systemic Lupus Erythematosus, a prevalent autoimmune disease. Through this study, up to forty SNPs and twenty genes were found to correlate to SLE, each having their specific function. It was found that certain genes, such as BLK, BANK1, and STAT4 were connected through co-expression, or transcription factors. It was also observed that SNPs were primarily intronic or intergenic, asserting that they influenced the expression of SLE through the control of gene expression. A gene map was also created, in order to bridge the gap between MHC class II genes, and SLE-related genes.

This study showed the assortment of SNPs that could affect SLE susceptibility, connection between gene function, as well as the linkage between the genes involved. Further experiments will provide information regarding the mechanisms by which genes regulate cytokines and symptoms of SLE.

6. Bibliography

1. Banchereau, J. & Pascual, V. Type I interferon in systemic lupus erythematosus and other autoimmune diseases. Immunity 25, 383?392 (2006).
2. Centers for Disease Control and Prevention. "Systemic Lupus Erythematosus (SLE)." *CDC*, www.cdc.gov/lupus/facts/detailed.html. Accessed 16 Aug. 2022.
3. Choi, Jinyoung et al. "The pathogenesis of systemic lupus erythematosus-an update." Current opinion in immunology vol. 24,6 (2012): 651-7. doi:10.1016/j.coi.2012.10.004
4. Dam, Elizabeth M et al. "The BANK1 SLE-risk variants are associated with alterations in peripheral B cell signaling and development in humans." Clinical immunology (Orlando, Fla.) vol. 173 (2016): 171-180. doi:10.1016/j.clim.2016.10.018
5. Elkon, Keith B, and Vivian V Stone. "Type I interferon and systemic lupus erythematosus." Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research vol. 31,11 (2011): 803-12. doi:10.1089/jir.2011.0045
6. Fava, Andrea, and Michelle Petri. "Systemic lupus erythematosus: Diagnosis and clinical management." *Journal of autoimmunity* vol. 96 (2019): 1-13. doi:10.1016/j.jaut.2018.11.001
7. Hussain, Nageen et al. "HLA association in SLE patients from Lahore-Pakistan." Bosnian journal of basic medical sciences vol. 11,1 (2011): 20-6. doi:10.17305/bjbms.2011.2618
8. J, De Azevêdo Silva et al. "Systemic Lupus Erythematosus: Old and New Susceptibility Genes versus Clinical Manifestations." Current genomics vol. 15,1 (2014): 52-65. doi:10.2174/1389202915011403061137 15
9. Jiang, S.H., Athanasopoulos, V., Ellyard, J.I. et al. Functional rare and low frequency variants in BLK and BANK1 contribute to human lupus. Nat Commun 10, 2201 (2019). https://doi.org/10.1038/s41467-019-10242-9

10. Ramos, Paula S et al. "Genetic factors predisposing to systemic lupus erythematosus and lupus nephritis." *Seminars in nephrology* vol. 30,2 (2010): 164-76. doi:10.1016/j.semnephrol.2010.01.007

11. Savitsky, David A et al. "Contribution of IRF5 in B cells to the development of murine SLE-like disease through its transcriptional control of the IgG2a locus." Proceedings of the National Academy of Sciences of the United States of America vol. 107,22 (2010): 10154-9. doi:10.1073/pnas.1005599107

12. Tokano, Y et al. "Levels of IL-12 in the sera of patients with systemic lupus erythematosus (SLE)--relation to Th1- and Th2-derived cytokines." Clinical and experimental immunology vol. 116,1 (1999): 169-73. doi:10.1046/j.1365-2249.1999.00862.x

13. Yang, H., Wang, K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. Nat Protoc 10, 1556?1566 (2015). https://doi.org/10.1038/nprot.2015.105