IJHSR
ijhighschoolresearch.org
REVIEW ARTICLE

Analysis of CD16$^{+/-}$ Monocytes and CD4^{+} T Cells to Identify Novel Gene Signatures and Develop a Diagnostic Tool for SLE

Maya LeBlanc

Abbey Park High School, 1455 Glen Abbey Gate, Oakville, Ontario, L6M 2G5, Canada; stemgirl25@gmail.com

ABSTRACT: Systemic lupus erythematosus (SLE) is an incurable chronic autoimmune disease that causes widespread inflammation and organ damage. Due to the lack of a single test to diagnose SLE, doctors use multiple general methods to diagnose the disease. This study evaluated gene expression in CD16^{+} monocytes, CD16^{-} monocytes, and CD4^{+} T lymphocyte cells to identify signatures unique to SLE, to improve diagnostic processes. Gene expression profiles from individuals diagnosed with SLE and females aged 24-29 controls were obtained from the Gene Expression Omnibus, and 54,675 gene probes were compared between healthy and SLE patients. The top five gene probes with increased differential expression between healthy and SLE patients were associated with the *ATP6V0C, UBA1, TGFB1, STAT1*, and *NFYC* genes. Quantile-quantile plots confirmed statistical appropriateness for genetic analysis. Further evaluation determined that the *ATP6V0C, UBA1, STAT1, NFYC*, and *TGFB1* genes associated with the CD16^{-} monocyte cell type represent a novel gene expression signature for SLE identification. Gene expression ranges were established for these probes, serving as a diagnostic tool for SLE. This tool can detect SLE in a single blood sample, which may improve diagnostic outcomes and reduce healthcare costs.

KEYWORDS: Biomedical and Health Sciences; Genetics and Molecular Biology of Disease; Systemic Lupus Erythematosus; Gene Expression Signatures; Transcriptomic Biomarkers.

Introduction

Systemic lupus erythematosus (SLE) is an incurable autoimmune disease in which the immune system generates antibodies that attack the body's own tissues, causing widespread inflammation and tissue damage.[1] SLE can affect the joints, skin, brain, lungs, kidneys, and blood vessels. Poor access to health care, late diagnosis, poor effectiveness of treatments, and imperfect adherence to therapeutic regimens may increase the damaging effects of SLE, resulting in complications and an increased risk of death. Between 2010 and 2016, the average number of deaths per year of US residents where SLE was identified as the underlying cause of death was 1,176. During the same 7-year period, SLE was recognized as a contributing cause of death in an average of 2,061 deaths per year.[1]

On average, it takes almost six years for people with SLE to be diagnosed with the disease from the time they first notice their symptoms.[2] SLE is known as "the great imitator" because its symptoms mimic many other illnesses. SLE symptoms can also be unclear, come and go, and change over time. In addition, SLE is diagnosed far more frequently in females than in males, a pattern that reflects a well-documented sex bias in both clinical presentation and existing datasets.

Furthermore, based on current resources and findings, it is important to note that the diagnosis of SLE currently cannot solely rely on a single test. Instead, doctors apply various methods to discover the presence of the disease. A thorough examination of the patient's medical history focuses on any genetic occurrences of SLE or other autoimmune disorders. Additionally, a comprehensive physical examination is performed to identify potential indicators such as skin rashes or other abnormal signs associated with SLE.[3]

Blood and urine tests, specifically the antinuclear antibody (ANA) test, are commonly used to assess the likelihood of the patient's immune system producing autoantibodies associated with SLE. While a positive ANA test is common among SLE patients, it does not confirm a SLE diagnosis conclusively. In the event of a positive ANA test result, the doctor typically orders further tests for antibodies specific to SLE.[3]

A skin or kidney biopsy, involving the removal of a tissue sample for microscopic examination, is sometimes recommended to detect potential signs of an autoimmune disease. However, knowing that none of these methods definitively determines SLE is crucial. Instead, their primary utility lies in excluding other conditions that may be mistaken for SLE. Recognizing that these diagnostic approaches do not offer a conclusive verdict on whether an individual has SLE is essential. Instead, they assist healthcare professionals in eliminating potential misdiagnoses and narrowing down the possibilities. Therefore, the great complexity of SLE diagnosis emphasizes the need for ongoing medical evaluation and collaboration between patients, healthcare providers, and medical researchers. SLE and other autoimmune disorders tend to run in families, but the inheritance pattern is unclear. People may inherit a gene variation that increases or decreases the risk of SLE, but in most cases, do not inherit the condition.[2] If a robust gene expression signature is identified for patients with SLE or at risk of SLE, this can be used as a diagnostic tool to improve diagnostic patient outcomes. A diagnostic tool that can conclusively identify SLE in patients would be novel and signi-

DOI: 10.36838/v7i9.80

ficantly impact patients and the health care system. Patients would have earlier access to treatments to help control symptoms and prevent further health decline. Healthcare costs to identify SLE would be substantially lower since many of the current testing regimes would not be necessary.

This investigation distinguishes itself from prior studies by focusing on the gene expression profiles of $CD16^-$ monocytes in systemic lupus erythematosus (SLE), a cell type subset that has been less explored in the context of SLE diagnostics. While previous research has highlighted the proinflammatory role of $CD16^+$ monocytes in SLE pathogenesis, particularly their involvement in T-cell activation and B-cell differentiation,[9] this study uniquely identifies a gene expression signature in $CD16^-$ monocytes.

$CD4^+$ T cells play a central role in coordinating the adaptive immune response. They act as "helper" cells that activate and direct other immune cells, including B cells, cytotoxic T cells, and macrophages, by secreting cytokines. In the context of systemic lupus erythematosus (SLE), $CD4^+$ T cells are critically involved in the loss of immune tolerance, which leads to the production of autoantibodies. Dysregulated $CD4^+$ T cells in SLE patients often exhibit abnormal activation, impaired regulatory function, and excessive help to B cells, leading to the formation of pathogenic autoantibody-producing plasma cells. Studies have also shown that $CD4^+$ T cells in SLE patients exhibit altered gene expression patterns linked to interferon signaling and pro-inflammatory cytokine production, further contributing to tissue damage and systemic inflammation.[17]

$CD16^+$ monocytes, also known as non-classical monocytes, are a subset of circulating monocytes that exhibit pro-inflammatory properties and are involved in patrolling the endothelium. In SLE, $CD16^+$ monocytes are found in elevated numbers and have been implicated in tissue infiltration and inflammation. These cells express higher levels of inflammatory cytokines such as TNF-α and IL-1β and contribute to the dysregulation of immune responses seen in SLE.[18]

In contrast, $CD16^-$ monocytes, or classical monocytes, are primarily involved in phagocytosis, the process by which a phagocyte (a type of white blood cell) surrounds and destroys foreign substances (such as bacteria) and removes dead cells.[10] These classical monocytes respond to infection and injury. While often considered less inflammatory, $CD16^-$ monocytes are important for understanding early immune activation and homeostasis. Interestingly, recent studies suggest that transcriptional reprogramming in these classical monocytes may occur early in SLE pathogenesis, even before overt clinical symptoms, making them valuable targets for early diagnosis and biomarker discovery.[19]

The purpose of this study was to evaluate gene expression in $CD16^+$ monocytes, $CD16^-$ monocytes, and $CD4^+$ T lymphocyte cells to identify gene expression signatures unique to systemic lupus erythematosus (SLE), which may offer an approach to improve diagnostic processes and outcomes for patients with SLE or at risk of acquiring SLE. Here it was investigated whether gene expression differed in $CD4^+$ T cells, $CD16^{+/-}$ monocytes between individuals with and without SLE.

Methods

Determine Gene Probe Values:

Initial research for the project included evaluating available gene expression data from various open-source repositories. The raw gene probe data chosen for this project were obtained from the Gene Expression Omnibus (GEO) hosted by the US National Center for Biotechnology Information (NCBI). GEO is a public genomics repository meant as an open source for scientific research. The datasets obtained from the GEO repository and used in this project were records GDS4888 ($CD4^+$ T lymphocytes), GDS4889 ($CD16^-$ monocytes), and GDS4890 ($CD16^+$ monocytes).

Gene expression differences between people with and without SLE were determined using microarray analysis. A total of 50 mL of peripheral blood was collected from each person. For $CD4^+$ T lymphocyte cells, this included six people with SLE (average age: 29.0±7.6) and four healthy people (average age: 24.8±0.5). CD16– monocyte cells included four people with SLE (average age: 26.5±1.7) and four healthy people (average age: 24.8±0.5). For $CD16^+$ monocyte cells, this included four people with SLE (average age: 26.5±1.7) and three healthy people (average age: 24.7±0.6). All people in the study were female. Erythrocytes were lysed in EL buffer, and then granulocytes were depleted using CD15-conjugated microbeads. The CD15-depleted fraction was stained with a CD14-fluorescein isothiocyanate antibody. Using a FACSAria cell sorter, the $CD4^+$ T cells, CD16- monocytes, and $CD16^+$ monocytes were isolated. After sorting, the cells were lysed with RLT buffer and frozen at -70°C. Total RNA was then isolated using an RNeasy mini kit. The generation of cRNA was accomplished by sample hybridization using HG-U133 Plus 2.0 arrays and scanning. The clinical characteristics of SLE and healthy persons are summarized in Table 1 below.

Table 1: Clinical characteristics of the female study participants, including CD^{4+} T lymphocyte cells from six SLE patients and four healthy individuals, CD^{16^-} monocyte cells from four SLE patients and four healthy individuals, and CD^{16+} monocyte cells from four SLE patients and three healthy individuals.

SLE / ND[a]	ID#	Collected cell type	Age	Sex	Disease activity: SLEDAI	ANA[b]	Anti-dsDNA[c]	Therapy
SLE	2 / M1	CD4+ T / CD16- Mo[d], CD16+ Mo	27	f	8	1:10240	48	MMF[e] 2000 mg/day
SLE	4 / M4	CD4+ T / CD16- Mo, CD16+ Mo	24	f	22	1:640	39	Pred.[f] 7 mg/day, HQ[g] 200mg/day, MMF 2000 mg/day
SLE	7	CD4+ T	42	f	6	1:5120	89	Pred. 10 mg/day, HQ 300 mg/day
SLE	8	CD4+ T	22	f	8	1:10240	1542	None
SLE	9	CD4+ T	34	f	8	1:2560	39	None
SLE	12	CD4+ T	25	f	10	1:5120	23	Pred. 10 mg/day, HQ 300 mg/day, CYC[h] 800 mg/month
SLE	M2	CD16- Mo, CD16+ Mo	28	f	2	1:2560	73	None
SLE	M3	CD16- Mo, CD16+ Mo	27	f	16	1:2560	130	None
ND	54	CD4+ T, CD16- Mo	25	f				
ND	55	CD4+ T, CD16- Mo, CD16+ Mo	24	f				
ND	56	CD4+ T, CD16- Mo, CD16+ Mo	25	f				
ND	57	CD4+ T, CD16- Mo, CD16+ Mo	25	f				

[a]ND: healthy donor.
[b]ANA: anti-nuclear antibody with cutoff for ANA titer <1:160.
[c]dsDNA-AK (U/ml) with cutoff 20 U/mL.
[d]Mo: monocytes.
[e]MMF: mycophenolate mofetil.
[f]Pred.: predonisolone.
[g]HQ: Hydroxychloroquine.
[h]CYC: cyclophosphamide.

The microarray data were analyzed through a multi-step process. First, data normalization and the generation of cell files were conducted using Affymetrix GCOS software. These cell files were then analyzed using the BioRetis database to perform group-wise comparisons and to filter for differentially expressed probe sets. To identify interferon (IFN)-regulated transcripts, the differentially expressed probe sets were compared with published reference lists. Finally, hierarchical cluster analysis was carried out using Genesis version 1.7.5. This comprehensive analysis produced gene probe expression values

DOI: 10.36838/v7i9.80

that were subsequently used for further investigation in the study.[16]

Identify Statistically Significant Gene Probes:

Gene probe expression differences were evaluated between SLE and healthy samples using a t-test. P-values of <0.0001 were identified as statistically significant, indicating a meaningful difference in gene expression between the two groups. Statistical comparison of the differentially expressed genes for the healthy versus the SLE cohort was performed using a one-tailed t-test to determine the resultant p-value. The value considered to be a statistically significant difference is a p-value less than 0.05. Variation in the data, which can affect the p-value calculation, was dealt with by using either a heteroscedastic or homoscedastic t-test. Before choosing the appropriate t-test, the variance of the data was calculated. The standard threshold for choosing a heteroscedastic t-test is 1.5 or greater.[4] This meant that a one-sided t-test was more appropriate. All the p-values were recalculated using a one-sided t-test. The process of performing the one-sided t-test and calculating the p-values for differentially expressed genes involved several key steps, as outlined below.

The datasets labeled “GDS4888”, “GDS4889”, and “GDS4890” were accessed through the website https://www.ncbi.nlm.nih.gov/, each representing a different cell type. In the "Downloads" box on the right side of the screen, the link to download the entire SOFT file was clicked for each of the three datasets. Each SOFT file was then opened in Excel after being downloaded. A new column titled “Variance Lupus” was created in the Excel spreadsheet. The following equation was inserted in the cell below: =VAR.S(C4:H4). This calculates the variance of the SLE data for that gene probe. A column labeled "Variance Healthy" was added beside "Variance Lupus," with the equation =VAR.S(G4:I4) inserted for variance calculation. After "Variance Healthy," a column labeled "Ratio" was created with the following equation: =IF(N4>O4, N4/O4, O4/N4). This equation ensures that if the Variance Lupus is greater than the Variance Healthy, the value for Variance Lupus is divided by the Variance Healthy; if the opposite is true, Variance Healthy is divided by Variance Lupus. The output represents the ratio between the variances of the Healthy and Lupus data. Next, a column titled "Homoscedastic T-test" was added to conduct the t-test if the variance was less than 1.5. The equation =TTEST(C4:H4, I4:L4,1,2) was used to compare the data from C4 to H4 with that from I4 to L4, with "1" indicating a one-sided t-test and "2" specifying a homoscedastic t-test. A column labeled "Heteroscedastic T-test" was created for use if the variance exceeded 1.5. This t-test equation was similar to that for the homoscedastic t-test, with the last number changed to "3" to denote a heteroscedastic t-test. Another column, titled "P Value Actual," was created with the equation =IF(T4>1.5, W4, V4). If the ratio was greater than 1.5, the heteroscedastic p-value was selected; otherwise, the homoscedastic p-value was chosen. Steps #1 to #11 were repeated for each gene probe in each dataset, with each dataset preferably placed on separate spreadsheets. Each spreadsheet contained over 54,000 rows.

Following this, the gene probes with the lowest p-values were identified by extracting all p-values, ID_REFs, and IDENTIFIERs for each gene probe across the datasets into a new Excel spreadsheet for comparison. Using Excel’s Sort function, p-values were sorted from smallest to largest to identify the smallest values. A new tab was created to filter for gene probes with p-values less than 0.001, and duplicates were removed using a formula. A combined list of unique gene probes was created, and VLOOKUP was used to match p-values for each gene probe across the datasets. A "Code" column was added to identify probes with p-values below 0.001. Data from this step were transferred to a new tab where redundant values were removed and gene probes were ranked based on their p-value significance. The final dataset included the top 5 gene probes with the lowest p-values, which were ranked and organized based on their average p-values across the datasets. Line graphs were then created for each of these top gene probes, comparing probe values between the healthy and SLE groups across the datasets.

Determination of Novel Gene Expression Signatures:

The top five gene probes were evaluated to identify those associated with vital functions, based on findings from current genetic research. These probes were selected for their critical roles in maintaining essential cellular and physiological processes. The next step involved defining a novel gene expression signature by identifying the cell type(s) with the lowest p-values for each selected gene probe. A gene expression signature refers to a specific gene, or a set of genes, that shows a strong statistical association with Systemic Lupus Erythematosus (SLE) and is linked to vital cellular functions within specific immune cell types. This signature is considered novel if it has not been previously described in scientific literature and demonstrates unique or previously unreported associations with SLE. To confirm novelty, the expression patterns of the identified genes were compared to existing publications, ensuring that the signature represents a new contribution to the understanding of SLE pathogenesis.

The three statistical tools used for analysis were QQ plots, histograms, and graphical cohort comparison. QQ plots assessed the distribution of p-values against a theoretical distribution, histograms visualized p-value distribution across cell types, and graphical analysis compared the differences between healthy and SLE cohorts.

The procedure used to create Quantile-Quantile (QQ) plots began with downloading and installing Python 3.12.2 (64-bit) from the official Python website (https://www.python.org/). For reference, the website https://support.minitab.com/en-us/minitab/21/integration/python-integration-guide/example-qq-plot/ was opened and left on-screen, as it provided guidance on how Minitab interfaces with Python to generate QQ plots. However, an alternative method proved more effective for this study. Visual Studio Code was downloaded from https://visualstudio.microsoft.com/ and installed following the latest instructions on the site. Next, Command Prompt was opened, and the command pip install mtbpy numpy matplotlib was entered to install the necessary Python module packages. Minitab Statistical Software was then downloaded and insta-

lled via a free trial from https://www.minitab.com/en-us/products/minitab/free-trial/, with care taken to use the latest version and follow the installation instructions. Minitab was pinned to the laptop, and the desktop version was launched. A zip file from the earlier support site was downloaded and unzipped in a designated folder. The file qq_plot.py from this archive was opened in Visual Studio Code and served as the bridge for Python-Minitab integration. In Minitab, the "Open" option was used to load the "Hospital test runs" data file, although its content was immediately deleted from the worksheet to prepare for new input. Relevant data was copied and pasted below the worksheet, and the Command Line was used to run the script with the customizable command PYSC "qq_plot.py" "Hospital A" "Hospital B". After clicking the "Run" button, QQ plots were generated and displayed. These plots could be copied by right-clicking and selecting "Copy Image." Additional functions were available through this interface. The Minitab-Python Interface code is given as a supplementary file to this journal. In addition to QQ plots, histograms were used to determine the distribution of p-values across the three immune cell types. Graphical comparisons between healthy and SLE cohorts were also conducted to assess whether the data differences trended positively or negatively.

Develop Gene Probe Expression Ranges for Diagnostic Tool:

This section describes the methods to identify the diagnostic tool's most significant gene probes and associated cell types. It also describes the methods used to develop gene probe expression ranges for the most important gene probes that would result in a p-value less than $1x10^{-4}$. These gene probe expression ranges have the potential to serve as a novel diagnostic tool for assessing individuals for the presence of SLE, or potentially identifying a genetic predisposition to developing SLE in the future. However, it is important to note that this application remains hypothetical and would require extensive clinical validation and large-scale studies before it could be implemented in practice.

The raw and p-value data for the 4889 (CD16-) datasets were opened in an Excel spreadsheet, and the gene probes used in the diagnostic tool were identified. Data for these probes were copied into a new spreadsheet, where a "Value" column was added. The p-value equation was modified to incorporate the "Value" column in the SLE portion of the p-value calculation. This process involved incrementally increasing the value in the "Value" column and generating a new p-value for each SLE value. The gene probe data, Value quantity, and p-value were then transferred to a Data Chart spreadsheet, where a "Criteria for P-Value" column was added with a value of 0.001, and a "Code" column was created with an equation to identify probes with p-values below 0.001. A line chart was created with gene probe values on the x-axis and a maximum p-value of 0.003. The range of gene probe values resulting in a p-value of 0.001 was considered statistically significant, and these steps were repeated for all evaluated gene probes.

In conclusion, regarding all methodology, the data obtained from the Gene Expression Omnibus (GEO) represented a full gene expression profile (54,675 probes) for healthy persons and persons with systemic lupus erythematosus (SLE). The data obtained measured differences in expression in 3 different cell types, which were $CD16^+$ monocyte, $CD16^-$ monocyte, and $CD4^+$ T lymphocyte cells for each cohort. T-test p-values (p-value) were calculated using Microsoft Excel for all gene probes in all three cell types. Variation in the data was accounted for by choosing either a heteroscedastic or homoscedastic t-test based on a variance test result threshold of 1.5. The variance calculation was performed using Microsoft Excel.

If interested in an in-depth methodology, a supplemental file is added to this journal.

Results and Discussion

Determine Gene Probe Values:

Due to the large amount of data generated during this study, including the raw and p-value data tables in this report was impossible. The raw data tables, including associated p-values, were labeled Table 2A, Table 2B, and Table 2C, one for each cell type. Due to the large data tables, it was impossible to include them in this publication.

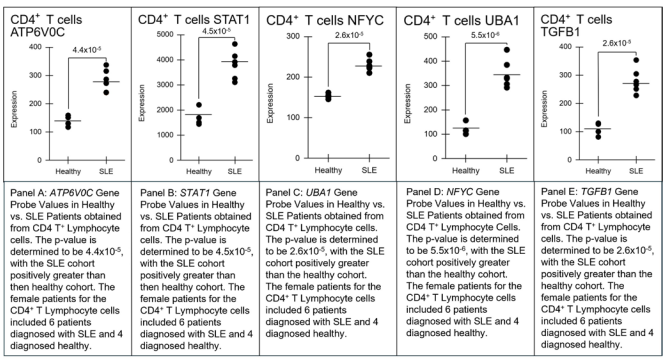
Identify Statistically Significant Gene Probes

The most significant differences between the SLE and healthy cohort gene expression across monocyte and CD4+ T cell subtypes were identified for the genes A*TP6V0C, TGFB1, STAT1, NFYC*, and *UBA1*. The lowest p-values for these genes ranged from $1.9x10^{-4}$ to $7.3x10^{-7}$. These p-values are very low and demonstrate that the difference in gene expression between the SLE and healthy cohorts is statistically significant.

Table 2: Lowest p values by cell type across different cell types, with each gene linked to its respective gene probe identifier (ID_REF). The table highlights the comparison between healthy individuals and those with SLE, revealing five gene probes with p-values ranging from 1.9×10^{-4} to 7.3×10^{-7}, indicating statistically significant differences in gene expression.

ID_REF	IDENTIFIER	p-Value		
(Gene Probe)	(Gene)	4888 (CD4+T)	4889 (CD16-)	4890 (CD16+)
200954_at	ATP6V0C	4.4E-05	1.4E-06	1.9E-04
200964_at	UBA1	2.6E-05	4.6E-06	3.8E-04
203085_s_at	TGFB1	2.6E-05	7.3E-07	4.3E-04
202215_s_at	NFYC	5.5E-06	2.0E-05	2.7E-04
200887_s_at	STAT1	4.5E-05	3.7E-04	3.9E-05

Scatterplots showing probe values for the diagnosed healthy and SLE cohorts, including associated p-values, for each of the top five gene probes and the three cell types were prepared. See Figures 1 to 3 below.



Panel A: *ATP6V0C* Gene Probe Values in Healthy vs. SLE Patients obtained from CD4 T+ Lymphocyte cells. The p-value is determined to be $4.4x10^{-5}$, with the SLE cohort positively greater than then healthy cohort. The female patients for the CD4+ T Lymphocyte cells included 6 patients diagnosed with SLE and 4 diagnosed healthy.

Panel B: *STAT1* Gene Probe Values in Healthy vs. SLE Patients obtained from CD4 T+ Lymphocyte cells. The p-value is determined to be $4.5x10^{-5}$, with the SLE cohort positively greater than then healthy cohort. The female patients for the CD4+ T Lymphocyte cells included 6 patients diagnosed with SLE and 4 diagnosed healthy.

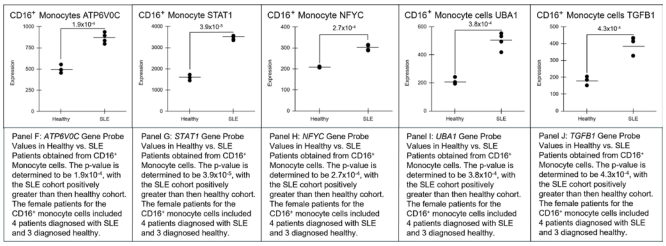
Panel C: *UBA1* Gene Probe Values in Healthy vs. SLE Patients obtained from CD4 T+ Lymphocyte cells. The p-value is determined to be $2.6x10^{-5}$, with the SLE cohort positively greater than the healthy cohort. The female patients for the CD4+ T Lymphocyte cells included 6 patients diagnosed with SLE and 4 diagnosed healthy.

Panel D: *NFYC* Gene Probe Values in Healthy vs. SLE Patients obtained from CD4 T+ Lymphocyte Cells. The p-value is determined to be $5.5x10^{-6}$, with the SLE cohort positively greater than the healthy cohort. The female patients for the CD4+ T Lymphocyte cells included 6 patients diagnosed with SLE and 4 diagnosed healthy.

Panel E: *TGFB1* Gene Probe Values in Healthy vs. SLE Patients obtained from CD4 T+ Lymphocyte cells. The p-value is determined to be $2.6x10^{-5}$, with the SLE cohort positively greater than the healthy cohort. The female patients for the CD4+ T Lymphocyte cells included 6 patients diagnosed with SLE and 4 diagnosed healthy.

Figure 1: The top five gene probes regarding the CD^4+ T Lymphocyte cell type, healthy vs. SLE patients gene expression.

DOI: 10.36838/v7i9.80



Panel F: *ATP6V0C* Gene Probe Values in Healthy vs. SLE Patients obtained from CD16+ Monocyte cells. The p-value is determined to be 1.9x10-4, with the SLE cohort positively greater than then healthy cohort. The female patients for the CD16+ monocyte cells included 4 patients diagnosed with SLE and 3 diagnosed healthy.

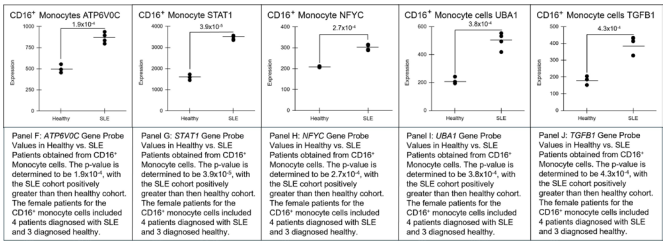
Panel G: *STAT1* Gene Probe Values in Healthy vs. SLE Patients obtained from CD16+ Monocyte cells. The p-value is determined to be 3.9x10-3, with the SLE cohort positively greater than then healthy cohort. The female patients for the CD16+ monocyte cells included 4 patients diagnosed with SLE and 3 diagnosed healthy.

Panel H: *NFYC* Gene Probe Values in Healthy vs. SLE Patients obtained from CD16+ Monocyte cells. The p-value is determined to be 2.7x10-4, with the SLE cohort positively greater than then healthy cohort. The female patients for the CD16+ monocyte cells included 4 patients diagnosed with SLE and 3 diagnosed healthy.

Panel I: *UBA1* Gene Probe Values in Healthy vs. SLE Patients obtained from CD16+ Monocyte cells. The p-value is determined to be 3.8x10-4, with the SLE cohort positively greater than then healthy cohort. The female patients for the CD16+ monocyte cells included 4 patients diagnosed with SLE and 3 diagnosed healthy.

Panel J: *TGFB1* Gene Probe Values in Healthy vs. SLE Patients obtained from CD16+ Monocyte cells. The p-value is determined to be 4.3x10-4, with the SLE cohort positively greater than then healthy cohort. The female patients for the CD16+ monocyte cells included 4 patients diagnosed with SLE and 3 diagnosed healthy.

Figure 2: The top five gene probes regarding the CD16+ Monocyte cell type healthy vs. SLE patients gene expression.



Panel F: *ATP6V0C* Gene Probe Values in Healthy vs. SLE Patients obtained from CD16+ Monocyte cells. The p-value is determined to be 1.9x10-4, with the SLE cohort positively greater than then healthy cohort. The female patients for the CD16+ monocyte cells included 4 patients diagnosed with SLE and 3 diagnosed healthy.

Panel G: *STAT1* Gene Probe Values in Healthy vs. SLE Patients obtained from CD16+ Monocyte cells. The p-value is determined to be 3.9x10-3, with the SLE cohort positively greater than then healthy cohort. The female patients for the CD16+ monocyte cells included 4 patients diagnosed with SLE and 3 diagnosed healthy.

Panel H: *NFYC* Gene Probe Values in Healthy vs. SLE Patients obtained from CD16+ Monocyte cells. The p-value is determined to be 2.7x10-4, with the SLE cohort positively greater than then healthy cohort. The female patients for the CD16+ monocyte cells included 4 patients diagnosed with SLE and 3 diagnosed healthy.

Panel I: *UBA1* Gene Probe Values in Healthy vs. SLE Patients obtained from CD16+ Monocyte cells. The p-value is determined to be 3.8x10-4, with the SLE cohort positively greater than then healthy cohort. The female patients for the CD16+ monocyte cells included 4 patients diagnosed with SLE and 3 diagnosed healthy.

Panel J: *TGFB1* Gene Probe Values in Healthy vs. SLE Patients obtained from CD16+ Monocyte cells. The p-value is determined to be 4.3x10-4, with the SLE cohort positively greater than then healthy cohort. The female patients for the CD16+ monocyte cells included 4 patients diagnosed with SLE and 3 diagnosed healthy.

Figure 3: The top five gene probes regarding the CD16- Monocyte cell type healthy vs. SLE patients gene expression.

Determination of Novel Gene Expression Signature

1) Identify Gene Probes Related to Vital Functions:

The top five gene probes with differential expression, *ATP6V0C, UBA1, TGFB1, NFYC*, and *STAT1*, were further evaluated to determine how they affect the human body and known interactions with other diseases.

ATP6V0C encodes a subunit of the vacuolar H+-ATPase (V-ATPase), a proton pump responsible for acidifying intracellular compartments such as lysosomes and endosomes. This gene plays a crucial role in maintaining cellular pH balance, which is essential for processes like endocytosis, protein degradation, and the proper functioning of organelles. The proper functioning of V-ATPase, and by extension *ATP6V0C*, is vital for cellular homeostasis and vesicular trafficking.[5]

UBA1 encodes the ubiquitin-activating enzyme 1, which is imperative for the ubiquitination process. *UBA1* activates ubiquitin molecules by attaching them to proteins, marking them for various fates, including degradation, alteration in activity, or changes in cellular location. This process is central to regulating cellular processes such as the cell cycle, stress responses, and protein turnover, ensuring that proteins are appropriately controlled within the cell.[7]

*TGFB*1 encodes Transforming Growth Factor Beta 1 (TGF-β1), a multifunctional cytokine that plays a pivotal role in regulating cell growth, differentiation, and immune function. TGF-β1 is involved in tissue repair and fibrosis by promoting extracellular matrix production and influencing immune responses. It also plays a significant role in immune suppression, helping regulate inflammation and maintain immune homeostasis, which is important for tissue homeostasis and wound healing.[12]

NFYC (Nuclear Transcription Factor Y Subunit C) is a transcriptional regulator that binds to the CAAT box sequence in the promoter regions of various eukaryotic genes. While it does not produce RNA directly, it assists in regulating transcription, the process of copying DNA into RNA, by RNA polymerase activity. This regulation can increase or decrease RNA transcript levels, thereby affecting the expression of target genes.[13]

STAT1 is another gene that encodes a transcription factor involved in the immune response, particularly in activating genes triggered by interferon, signaling proteins that help the body fight infections and cancer. Interferon is a natural substance produced in the body by white blood cells that help the body's immune system fight infection and other diseases, such as cancer.[6] The protein encoded by *STAT1* immune system is involved in transmitting signals within cells, particularly in response to interferons, which are signaling proteins that play a key role in the immune response to viral infections. When interferons bind to their receptors on the surface of a cell, they activate *STAT1* and cause it to move into the cell's nucleus.[14]

All five genes are on the list of priority gene probes for further evaluation.

2) Define Novel Gene Expression Signature:

The p-value data for the top three gene probes were evaluated to determine whether one or more cell types would be used as a diagnostic tool.

In this analysis, the CD16- monocyte cell type was selected not solely based on having the lowest p-values, but rather because the p-values reflected consistent and statistically robust differences across all three gene probes. While p-values indicate the likelihood that observed differences are not due to random variation, they do not capture the magnitude of expression change. To address this, complementary analyses assessing effect sizes (e.g., fold-change) could be performed to evaluate the biological relevance of the differences. However, to establish a statistically sound candidate cell type in this initial assessment, statistical significance was prioritized to ensure that any observed differences were reliable across all gene probes.

Further research identified a list of the current known genes associated with SLE. These genes are shown in Table 3. The three genes identified in this study, *ATP6V0C, UBA1, TGFB1, UBA1*, and *STAT1*, are not included in Table 3. Therefore, the *ATP6V0C, UBA1, TGFB1, UBA1*, and *STAT1* genes associated with the CD16- monocyte cell type represent a unique and novel gene expression signature for identifying SLE in patients.

Table 3: Genes Associated with SLE[8]. However, the three genes identified in this study, *ATP6V0C, UBA1, NFYC, STAT1*, and *TGFB1*, are not included, highlighting their unique and novel role in gene expression signatures for identifying SLE in patients with CD16- monocyte cell types.

Table 1: The top five gene probes regarding the CD16+ Monocyte cell type healthy vs. SLE patients gene expression.

Gene	Location	Odds ratio	Best *P* value	Population
PTPN22	1p13.2	1.4	3.4 × 10−12	EU, HA
FCGR2A, FCGR3B	1q23	0.74	6.8 × 10−7*a*	EU, AA, AS
NCF2	1q25	1.19	4.62 × 10−20	EU, AS
CRP	1q21	0.49	9.2 × 10−14	EU, AA
TNFSF4	1q25	1.46	2.5 × 10−32	EU, AS, HA
IL10	1q31–q32	1.19	4.0 × 10−8	EU, AA
Complement genes	1p36		Convincing	EU
RASGRP3	2p24.1	0.7	1.3 × 10−15	AS
IFIH1	2q24	1.11	1.6 × 10−8	EU

STAT4	2q32.2	1.55	5.17 × 10−42	EU, AA, AS, HA
PXK	3p14.3	1.25	7.1 × 10−9	EU
TREX1	3p21.31	44.65	8.5 × 10−11	EU
BANK1	4q24	1.31	2.62 × 10−13	EU, AA, AS, HA
IL2/IL21	4q26	1.16	2.2 × 10−8	EU, AA, AS
TNIP1	5q32	1.27	1.67 × 10−9	EU, AA, AS
Mir146a	6q5	1.29	2.74 × 20−8	AA, AS
HLA and other genes	6p21.3	2.35	1.27 × 10−51	EU, AS, AA, HA
ATG5	6q21	1.25	5.2 × 10−12	EU, AS
TNFAIP3	6q23	1.72	1.3 × 10−17	EU, AA, AS
IKZF1	7p13	0.72	2.8 × 10−23	AS, EU
JAZF1	7p15.2	1.19	1.5 × 10−9	EU
IRF5	7q32	1.54	3.611 × 10−19	EU, AA, AS, HA
XKR6	8p23.1	1.23	2.5 × 10−11	EU
BLK	8p23	0.69	2.1 × 10−24	EU, AA, AS, HA
LYN	8q13	0.77	5.4 × 10−9	EU, AA, AS
LRRC18, WDFY4	10q11.23	1.24	7.2 × 10−12	AS
CD44	11p13	0.71	4.0 × 10−12	EU, AS, AA
PHRF1/IRF7/KIAA1542	11p15.5	0.78	3 × 10−10	EU, AA
ETS1	11q24.3	1.37	1.8 × 10−25	AS
SLC15A4	12q24.32	1.26	1.77 × 10−11	AS
ELF1	13q13	1.26	1.5 × 10−8	AS
ITGAM	16p11.2	1.62	1.61 × 10−23	EU, AS, HA
PRKCB	16p11.2	0.81	1.4 × 10−9	AS
IRF8	16q24.1	1.16	2.3 × 10−9	EU
TYK2	19p13.2	1.2	3.88 × 10−8	EU
CD40	20q12	0.63	2.0 × 10−8	EU
UBE2L3	22q11.21	0.78	1.48 × 10−16	EU, AS
TLR7	Xp22.3	1.67	6.5 × 10−10	AS
IRAK1/MECP2	Xq37	1.39	6.65 × 10−11	EU, AS, HA

Table 3 above from the Journal of Leukocyte Biology (2012) presents a comprehensive overview of over 50 genes statistically associated with SLE. Each entry includes the gene's name, chromosomal location, odds ratio indicating the strength of association with SLE, p-value demonstrating statistical significance, and the specific populations in which the gene variant was studied: European (EU), African American (AA), Asian (AS), and Hispanic American (HA). An odds ratio greater than 1 reflects an increased risk of developing SLE, while values below 1 suggest a potential protective effect. The chart was selected for analysis to benchmark newly identified genes, *ATP6V0C, UBA1, STAT1, NFYC*, and *TGFB1*, against existing SLE-associated genes. The absence of these genes in Table 3 supports their novelty and strengthens the claim that they represent a unique gene expression signature. This comparison validates the identification of previously unreported genetic markers in CD16- monocyte cells as well.

3) Final Statistical Analysis:

A) QQ Plots:

When reporting probe values for genes, there are only examples of genes where the value is higher in SLE patients compared to healthy patients. Generating a Q-Q plot is a common way to showcase that the test has a proper significance distribution. The results of the QQ plots, as shown in Figures 4 to 11, determine that the data is statistically appropriate for analysis.

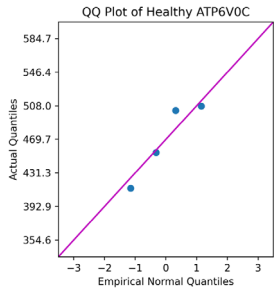


Figure 4: Quantile-Quantile (QQ) plot graph for gene probe values regarding the probe associated with the ATP6V0C gene, for the Healthy cohort, with respect to the CD16- monocyte cell type. cell type, healthy vs. SLE patients gene expression.

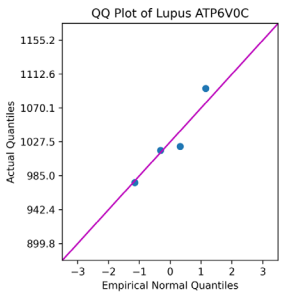


Figure 5: Quantile-Quantile (QQ) plot graph for gene probe values regarding the probe associated with the ATP6V0C gene, for the Healthy cohort, with respect to the CD16- monocyte cell type. cell type, healthy vs. SLE patients gene expression.

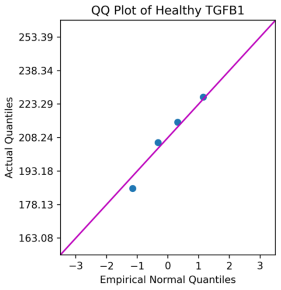


Figure 6: Quantile-Quantile plot graph for gene probe values regarding the probe associated with the *TGFB1* gene, for the Healthy cohort, with respect to the CD16- monocyte.

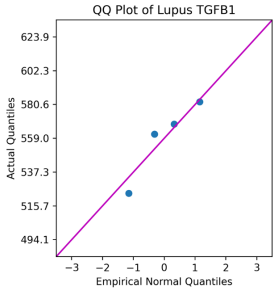


Figure 7: Quantile-Quantile plot graph for gene probe values regarding the probe associated with the *TGFB1* gene, for the SLE cohort, with respect to the CD16- monocyte.

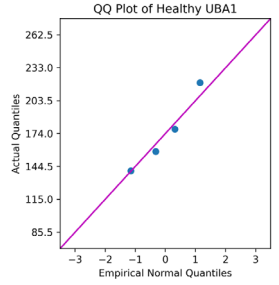


Figure 8: Quantile-Quantile plot graph for gene probe values regarding the probe associated with the UBA1 gene, for the Healthy cohort, with respect to the CD16- monocyte.

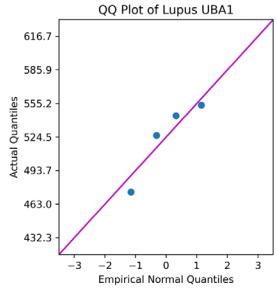


Figure 9: Quantile-Quantile plot graph for gene probe values regarding the probe associated with the UBA1 gene, for the SLE cohort, with respect to the CD16- monocyte.

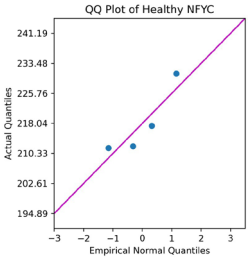


Figure 10: Quantile-Quantile (QQ) plot graph for gene probe values regarding the probe associated with the NFYC gene, for the Healthy cohort, with respect to the CD16- monocyte cell type.

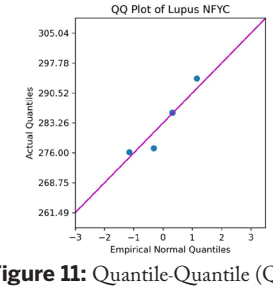


Figure 11: Quantile-Quantile (QQ) plot graph for gene probe values regarding the probe associated with the NFYC gene, for the SLE cohort, with respect to the CD16- monocyte cell type.

DOI: 10.36838/v7i9.80

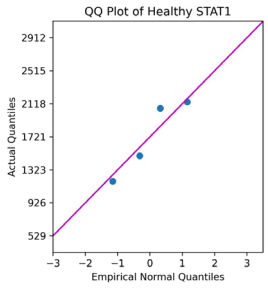


Figure 12: Quantile-Quantile (QQ) plot graph for gene probe values regarding the probe associated with the STAT1 gene, for the Healthy cohort, with respect to the CD16⁻ monocyte cell type.

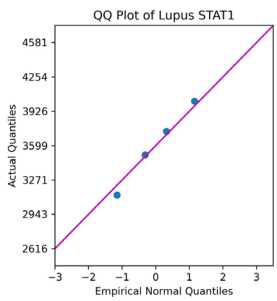


Figure 13: Quantile-Quantile (QQ) plot graph for gene probe values regarding the probe associated with the STAT1 gene, for the SLE cohort, with respect to the CD16⁻ monocyte cell type.

A standard test to determine if genetic data is statistically appropriate and follows a normal distribution is a test called a Quantile-Quantile (QQ) plot. A QQ plot plots a distribution's observed quantiles versus the ideal distribution. Where quantiles are regular, equally spaced intervals of a random variable divide the random variable into units of equal distribution.[11] Above are the QQ plots for the healthy and SLE cohorts associated with the top three gene probes evaluated.

B) Histograms:

Histograms were created to visualize the distribution of p-values derived from differential gene expression analyses between healthy controls and SLE patients. While p-values alone do not reflect the magnitude or biological relevance of gene regulation, they remain a useful tool for identifying statistically significant patterns within large-scale expression datasets. The raw gene probe data in Table 1 are shown in Figures 12, 13, and 14, as frequency histograms for each cell type, CD16- monocyte, CD16+ monocyte, and CD4+ T lymphocyte. These histograms show the number of times the p-values for a specific range of values occur in the dataset. These histograms show relatively high p-values in the very low p-value ranges. This is normal for gene probe data and indicates that there are outliers that have statistically significant p-values.

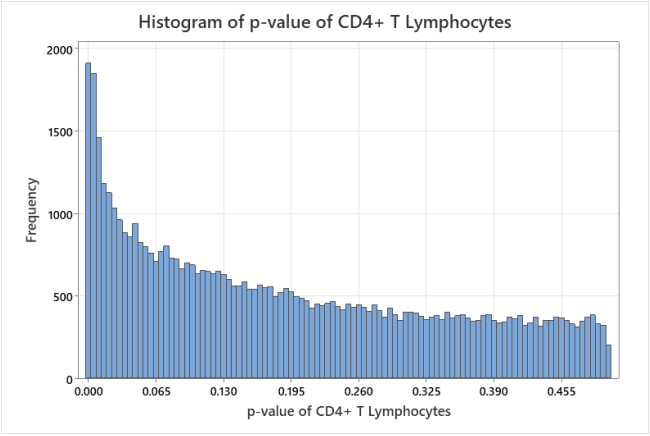


Figure 14: Histogram of p-values for CD4⁺ lymphocyte dataset. The dataset represented full gene probe analysis results (54,675 probes x 1 cell type CD⁴+ T Lymphocyte) for healthy persons and persons with systemic lupus erythematosus (SLE).

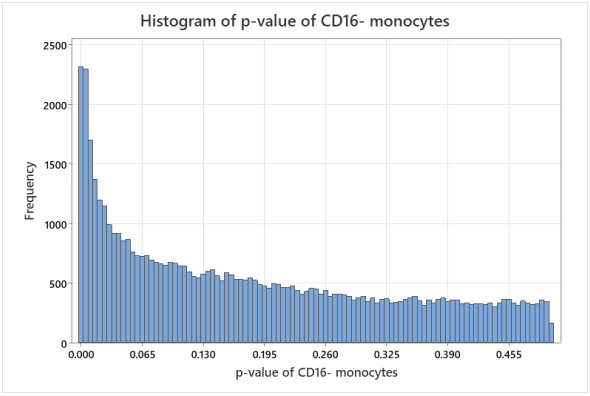


Figure 15: Histogram of p-values for CD16⁻ monocyte dataset. The dataset represented full gene probe analysis results (54,675 probes x 1 cell type CD16⁻ Monocyte) for healthy persons and persons with systemic lupus erythematosus (SLE).

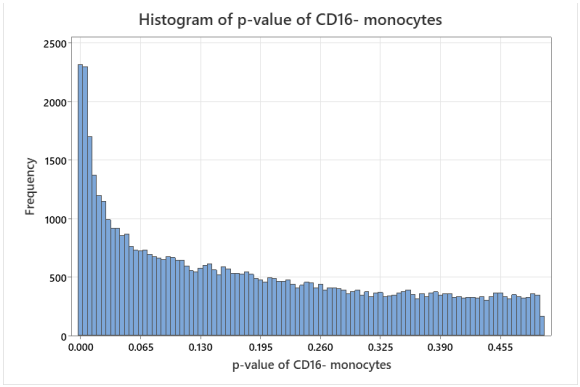


Figure 16: Histogram of p-values for CD16⁺ monocyte dataset. The dataset represented full gene probe analysis results (54,675 probes x 1 cell type CD16⁺ Monocyte) for healthy persons and persons with systemic lupus erythematosus (SLE).

C) Healthy versus SLE Charts:

It was noted previously that all the line graphs in Figures 1 to 3 showed gene probe data for the SLE cohort, which was higher in value than the healthy cohort. Two gene probe scatterplots, Figure 17, show that there are gene probes for which the healthy cohort has higher values than the SLE cohort. Many more gene probes with this trend demonstrate a variety in the data. Each bar is the expression value in an individual T cell.

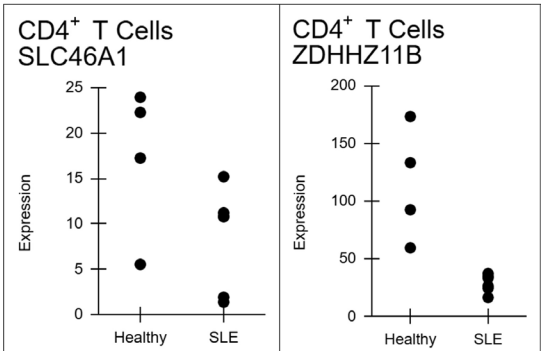


Figure 17: Two Gene Prove Values in Healthy vs. SLE Patients to Demonstrate Different Cohort Positioning.

4. Develop Gene Probe Expression Ranges for Diagnostic Tool:

Two gene probe values were identified, resulting in a p-value of $1x10^{-4}$ for each of the top three gene probes from the $CD16^-$ monocyte cell dataset. These two gene probe values represent the highest and lowest gene probe expressions, resulting in a statistically significant difference between the healthy and SLE cohorts, based on minimum p-value criteria of $1x10^{-4}$. Charts were prepared for each of the top three gene probes to show this data. See Figures 18 to 22.

To note, expression ranges are method-dependent, with different gene expression platforms (e.g., microarray, RNA-seq, qPCR) potentially producing varying absolute values. In this analysis, ranges were derived specifically based on the methodology employed to ensure internal consistency. The analysis was performed to provide a structured framework for interpreting expression differences and selecting the most suitable candidate cell type.

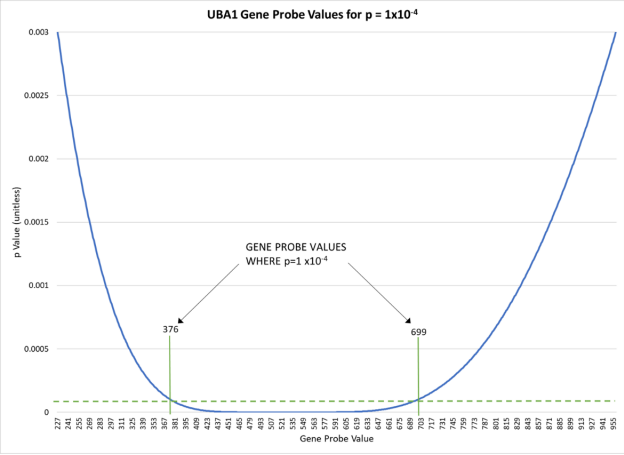


Figure 18: *UBA1* Gene Probe Values for p = $1x10^{-4}$. Displays the statistically significant UBA1 gene probe range (p<0.0001) for detecting the presence of SLE using $CD16^-$ monocyte cells. Specifically, the *UBA1* gene probe range is between 376 and 699.

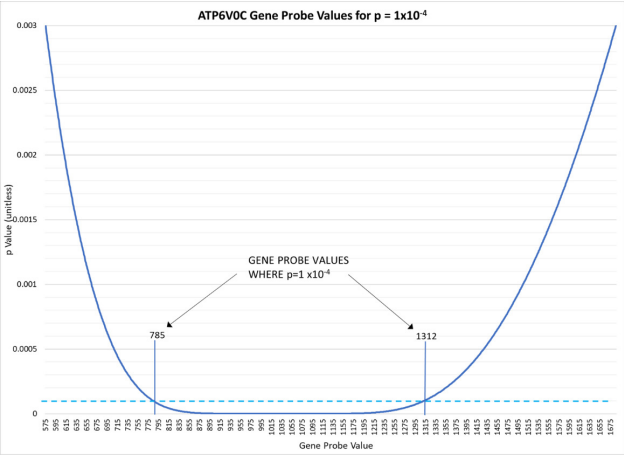


Figure 19: *ATP6V0C* Gene Probe Values for p = $1x10^{-4}$. Displays the statistically significant *ATP6V0C* gene probe range (p<0.0001) for detecting the presence of SLE using $CD16^-$ monocyte cells. Specifically, the *ATP6V0C* gene probe range is between 785 and 1312.

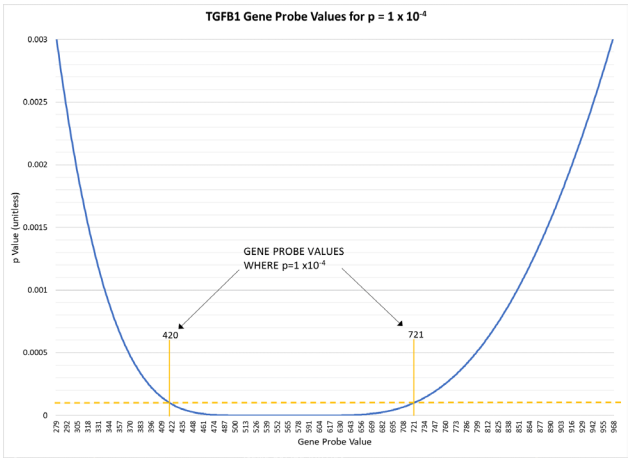


Figure 20: *TGFB1* Gene Probe Values for p = $1x10^{-4}$. This figure provides a comprehensive overview of the statistically significant TGFB1 gene probe range (p<0.0001) for detecting SLE through the analysis of $CD16^-$ monocyte cells, with gene probe values spanning from 420 to 721.

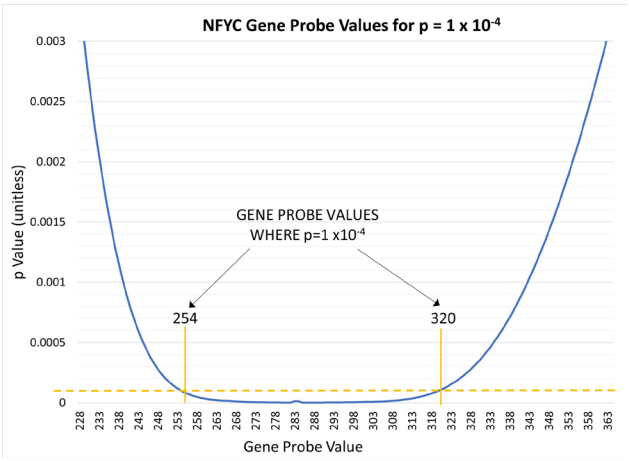


Figure 21: *NFYC* Gene Probe Values for p = $1x10^{-4}$. Displays the statistically significant NFYC gene probe range (p<0.0001) for detecting the presence of SLE using $CD16^-$ monocyte cells. Specifically, the *NFYC* gene probe range is between 254 and 320.

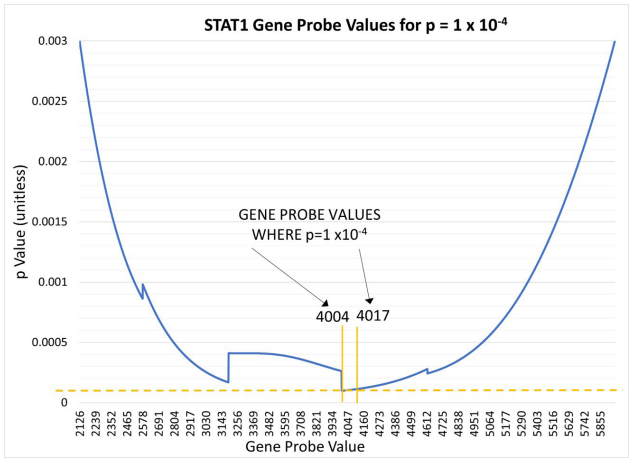


Figure 22: *STAT1* Gene Probe Values for p = $1x10^{-4}$. Displays the statistically significant STAT1 gene probe range (p<0.0001) for detecting the presence of SLE using $CD16^-$ monocyte cells. Specifically, the STAT1 gene probe range is between 4004 and 4017.

DOI: 10.36838/v7i9.80

The graph above shows "abnormality" in comparison to the other graphs due to the variance set to 1.5. Below is a graph for *STAT1* with Variance disregarded, and therefore, the homoscedastic t-test is only regarded as the final p-value to analyze. This will remove the "jagged edges" visually displayed.

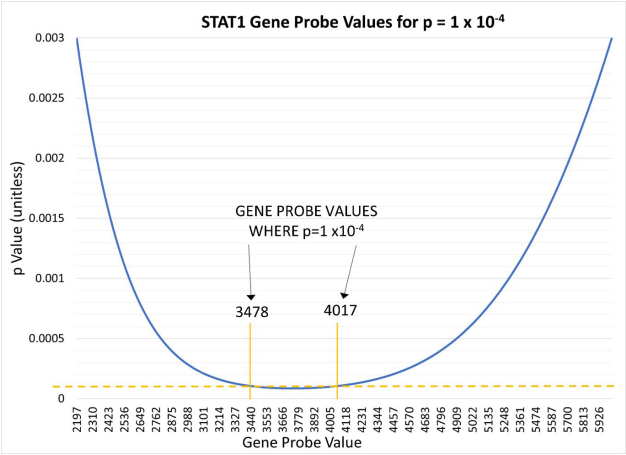


Figure 23: *STAT1* Gene Probe Values for p = 1x10^{-4}. P-values are produced through only the homoscedastic t-test. Displays the statistically significant STAT1 gene probe range (p<0.0001) for detecting the presence of SLE using CD16^{-} monocyte cells. Specifically, the STAT1 gene probe range is between 3478 and 4017.

Conclusion

The following are the significant conclusions from this study.

The study determined a statistical difference in the gene probe values for genes *ATP6V0C, UBA1, NFYC, STAT1*, and *TGFB1* between healthy people and people with SLE. The t-test p-values (p-value) from this analysis for these genes were less than 1 x 10-4 for all cell types in the study, CD16^{-} monocyte, CD16^{+} monocyte, and CD4^{+} T lymphocyte.

Furthermore, it was determined that the CD16- monocyte cell type is the best indicator of statistical difference in gene probe expression in this study for SLE in the *ATP6V0C, UBA1, NFYC, STAT1*, and *TGFB1* genes. The p-values from this analysis ranged between 1.4 x 10^{-6} to 7.3 x 10^{-7}.

The study results were compared to a list of known genes associated with SLE. It was determined that the gene probes in this study, related to the *ATP6V0C, UBA1, NFYC, STAT1*, and *TGFB1* genes and the CD16^{-} monocyte cell type, represent a novel gene expression signature for the identification of SLE.

The criteria for a novel diagnostic test method were developed to detect the presence of SLE in patients. The criteria are based on the high and low gene probe expression values identified in this study as statistically significant in SLE patients. Using these high and low gene probe values that define the gene probe expression ranges that are statistically significant in SLE patients a diagnostic test method could be developed to test patients for the presence of SLE. First, a blood sample will be obtained from a patient. The CD16^{-} monocyte cells would be isolated from the blood sample. Then, the CD16^{-} monocyte cells would be analyzed for the three gene probes defined in this study and associated with the *ATP6V0C, UBA1, NFYC, STAT1*, and *TGFB1* genes. This diagnostic tool would be a quick and relatively simple way to determine if a person has SLE. Introducing this innovative diagnostic method could transform the lives of SLE patients globally. It quickly and accurately detects SLE, so patients can receive timely treatment, vastly improving their quality of life and potentially saving lives. This test could ease the financial burden on healthcare systems by simplifying diagnosis, allowing resources to be redirected toward patient support and research into new SLE treatments and a possible cure.

The sample size for this study was relatively small. Further analysis with additional persons in the SLE and healthy cohorts would improve the study's statistical power and increase confidence in rejecting the null hypothesis. Further long-term research, including persons who are asymptomatic for SLE, is required to determine if this test can also determine whether a healthy person is genetically predisposed to SLE in the future. This is a very interesting study area since I am unaware of any quantitative methods for determining a person's predisposition to SLE.

Acknowledgments

I would like to greatly thank Dr. Mathieu Lupien, Ornela Kljakic, and Dr. Guillaume Bourque for their crucial guidance in defining the project's focus. Thank you, Rishi Singh and Minitab Statistical Software, for helping with the Python and Minitab Statistical Software Interface challenges. Thank you, Dr. Dawn Bowdish, Dr. Jessica A. Breznik, and Dr. Konstantinos Tselios, for serving as expert reviewers for this report and providing invaluable comments and insights. I would like to finally express my gratitude to my mother and father and thank those who have dedicated their lives to researching, understanding, and improving systemic lupus erythematosus and diagnostics. I wouldn't be here without them.

All glory be to God.

References

1. Centers for Disease Control and Prevention. Diagnosing and Treating Lupus | CDC. www.cdc.gov. https://www.cdc.gov/lupus/basics/diagnosing.htm.
2. Lupus facts and statistics. (2021, July 23). Lupus Foundation of America. https://www.lupus.org/resources/lupus-facts-and-statistics
3. CDC. Systemic Lupus Erythematosus (SLE) | CDC. www.cdc.gov. https://www.cdc.gov/lupus/facts/detailed.html#:~:text=doing%20about%20SLE%3F-.
4. Stephanie. Homoscedasticity / Homogeneity of Variance/ Assumption of Equal Variance. Statistics How To. https://www.statisticshowto.com/homoscedasticity/.
5. ATP6V0C ATPase H+ transporting V0 subunit c [Homo sapiens (human)] - Gene - NCBI. (2025). Nih.gov. https://www.ncbi.nlm.nih.gov/gene/527
6. NCI Dictionary of Cancer Terms. National Cancer Institute. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/interferon.
7. UBA1 gene: MedlinePlus Genetics. (2017). Medlineplus.gov. https://medlineplus.gov/genetics/gene/uba1/
8. Vaughn, S. E.; Kottyan, L. C.; Munroe, M. E.; Harley, J. B. Genetic Susceptibility to Lupus: The Biological Basis of Genetic Risk Found in B Cell Signaling Pathways. Journal of Leukocyte Biology 2012, 92 (3), 577–591. https://doi.org/10.1189/jlb.0212095.
9. Zhu, H., Hu, F., Sun, X., Zhang, X., Zhu, L., Liu, X., Li, X., Xu,

L., Shi, L., Gan, Y., & Su, Y. (2016). CD16+ Monocyte Subset Was Enriched and Functionally Exacerbated in Driving T-Cell Activation and B-Cell Response in Systemic Lupus Erythematosus. Frontiers in Immunology, 7. https://doi.org/10.3389/fimmu.2016.00512
10. National Cancer Institute. (2011, February 2). https://www.cancer.gov/publications/dictionaries/cancer-terms/def/phagocytosis. Www.cancer.gov. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/phagocytosis
11. Adkins, D. Introduction to Statistical Genetics. https://statisticalhorizons.com/wp-content/uploads/2022/04/SG-Sample-Materials-2.pdf.
12. Vivian Weiwen Xue, Jeff Yat-Fai Chung, Alexandra, C., Alvin Ho-Kwan Cheung, Kang, W., Eric W.-F. Lam, Kam Tong Leung, To, K., Lan, H., & Patrick Ming-Kuen Tang. (2020). Transforming Growth Factor-β: A Multifunctional Regulator of Cancer Immunity. Cancers, 12(11), 3099–3099. https://doi.org/10.3390/cancers12113099
13. NFYC protein expression summary - The Human Protein Atlas. (2021). Proteinatlas.org. https://www.proteinatlas.org/ENSG00000066136-NFYC
14. STAT1 gene: MedlinePlus Genetics. (n.d.). Medlineplus.gov. https://medlineplus.gov/genetics/gene/stat1/
15. Tin, A., Marten, J., Halperin Kuhns, V. L., Li, Y., Wuttke, M., Kirsten, H., Sieber, K. B., Qiu, C., Gorski, M., Yu, Z., Giri, A., Sveinbjornsson, G., Li, M., Chu, A. Y., Hoppmann, A., O'Connor, L. J., Prins, B., Nutile, T., Noce, D., & Akiyama, M. (2019). Target genes, variants, tissues and transcriptional pathways influencing human serum urate levels. Nature Genetics, 51(10), 1459–1474. https://doi.org/10.1038/s41588-019-0504-x
16. geo. (2025). GEO DataSet Browser. Nih.gov. https://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS4888#details
17. Gurkan, I., Ranganathan, A., Yang, X., Horton, W. E., Todman, M., Huckle, J., Pleshko, N., & Spencer, R. G. (2010). Modification of osteoarthritis in the guinea pig with pulsed low-intensity ultrasound treatment. Osteoarthritis and Cartilage, 18(5), 724–733. https://doi.org/10.1016/j.joca.2010.01.006
18. Wise, E. M., Henao, J. P., Gomez, H., Snyder, J., Roolf, P., & Orebaugh, S. L. (2015). The impact of a cadaver-based airway lab on critical care fellows' direct laryngoscopy skills. Anaesthesia and Intensive Care, 43(2), 224–229. https://doi.org/10.1177/0310057X1504300213
19. Retraction: "Concurrent inhibition of NF-κB, cyclooxygenase-2, and epidermal growth factor receptor leads to greater anti-tumor activity in pancreatic cancer" by Ali *et al.* (2016). Journal of Cellular Biochemistry, 117(8), 1961. https://doi.org/10.1002/jcb.25586

Author

Maya LeBlanc is a high school graduate in Ontario, Canada, eager to expand her knowledge in math, physics, computational/mathematical biology, and robotics. She plans to pursue her passion in mechanical/mechatronics engineering at university and become a well-rounded mechatronics engineer with applications in biomedical sciences and the aerospace industry.

DOI: 10.36838/v7i9.80