

## GOPEN ACCESS

**Citation:** Brant EJ, Rietman EA, Klement GL, Cavaglia M, Tuszynski JA (2020) Personalized therapy design for systemic *lupus erythematosus* based on the analysis of protein-protein interaction networks. PLoS ONE 15(3): e0226883. https://doi. org/10.1371/journal.pone.0226883

**Editor:** Xu-jie Zhou, Peking University First Hospital, CHINA

Received: March 6, 2019

Accepted: December 8, 2019

Published: March 19, 2020

**Copyright:** © 2020 Brant et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The input data underlying the results of this paper are publicly available here: (https://www.ncbi.nlm.nih.gov/geo/ query/acc.cgi?acc=GSE45219). Interested researchers can also request access to the result files by contacting the authors.

**Funding:** The authors received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

RESEARCH ARTICLE

# Personalized therapy design for systemic *lupus erythematosus* based on the analysis of protein-protein interaction networks

Elizabeth J. Brant<sup>1</sup>, Edward A. Rietman<sup>2,3</sup>, Giannoula Lakka Klement<sup>4</sup>, Marco Cavaglia<sup>5</sup>, Jack A. Tuszynski<sup>6,7</sup>\*

Nephrology, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, United States of America,
 BINDS lab, College of Information and Computer Sciences, University of Massachusetts Amherst,
 Amherst, Massachusetts, United States of America, 3 Department of Mechanical and Industrial Engineering,
 University of Mass, Amherst, Massachusetts, United States of America, 4 CSTS, Toronto, Ontario, Canada,
 ACTISMED, srl, Torino, Italy, 6 DIMEAS, Politecnico di Torino, Torino, Italy, 7 Department of Oncology,
 Cross Cancer Institute, University of Alberta, Edmonton, Alberta, Canada

\* jackt@ualberta.ca

## Abstract

We analyzed protein expression data for Lupus patients, which have been obtained from publicly available databases. A combination of systems biology and statistical thermodynamics approaches was used to extract topological properties of the associated protein-protein interaction networks for each of the 291 patients whose samples were used to provide the molecular data. We have concluded that among the many proteins that appear to play critical roles in this pathology, most of them are either ribosomal proteins, ubiquitination pathway proteins or heat shock proteins. We propose some of the proteins identified in this study to be considered for drug targeting.

### Introduction

Systemic lupus erythematosus (SLE) is a unique autoimmune disease with multiple pathologies including organ damage to kidney, skin, lungs, brain and heart, among others. Women of childbearing age and African-American persons are largely affected, with a ratio of 9:1 compared to general population. Its pathogenesis is not yet clearly defined but is generally thought to be due to a complex interplay between genetics [1-4], environmental and female sex hormone [5], and epigenetics [6]. Genome-wide association studies have identified 46 single nucleotide polymorphisms (SNPs) shown to predispose to SLE, 30 lie within noncoding regions of the human genome [7]. The most common genetic predisposition is found at the MHC locus and specially the HLA-DR2 and HLA-DR3, with a hazard ratio of 2. Genetically-predisposing variants involve some associated with innate immunity (IRF5, STAT4, IRAK1, TNFAIP3, SPP1, and TLR7), most of which are associated with interferon alpha pathways. Still other predisposing genes involve lymphocyte signaling (PTPN22, OX40L, PD-1, BANK-1, LYN, BLK), each of which plays a role in the activation or suppression of T cell or B cell activation or survival. In addition to genes, epigenetic modifications are important in the pathogenesis of SLE. These include hypomethylation of DNA, which influences transcription into protein. Genetic factors that confer the highest HR of 5 to 25, although

rare, are deficiencies of the complement components C1q (required to clear apoptotic cells), C4A and B, C2, or the presence of a mutated TREX1 gene. Genetic information accounts for only 18 percent of susceptibility to SLE, suggesting a large component of environmental or epigenetic influences [8]. Studies on the expression profiles of lncRNAs in T cells of SLE patients revealed some lncRNAs whose expressions might correlate with disease activity of SLE patients [9]. Histone H3 lysine 4 trimethylation (H3K4me3) is an important epigenetic modification, which is associated with active transcription and it has been shown that there are significant alterations of H3K4me3 in the peripheral blood mononuclear cells of SLE patients [10]. Other studies examined H3K4me3 breadth at transcription start sites (TSS) in primary monocytes and its association with differential gene transcription in SLE, providing evidence that TSS might be a crucial regulator responsible for transcription changes in SLE [11]. The data of individual T cell miRNA expression profiles in the literature for SLE risk or pathogenesis are quite variable [12, 13, 9]. The study in ref. [14] concluded that a number of elevated miRNAs could potentially become biomarkers for immunopathogenesis of SLE9. These biomarkers include elevation of miR-17-92 cluster, miR-21, miR-296, miR-126, miR-148a, miR-224, miR-524-5p, and suppression of miR-31, miR-125a, miR-125b, miR-142-3p, miR-142-5p and miR-146a. In addition, these biomarkers are found intriguingly correlated with T cell subset alteration, aberrant cytokine/chemokine release, altered gene transcription and immune cell signaling abnormalities in SLE [15].

Moreover, urinary exosomal miRNA profiling was also investigated in connection with biomarkers for lupus nephritis [16–17]. These include increased miR-125a, miR-146, miR-150 and miR-155, and decreased miR-141, miR-192 and miR-200a. When exploring the miRNA expression profiles in the damaged target tissues, the authors of ref. [18] directly identified, confirmed and explicated miR-30c-5p, miR-1273e and miR-3201 in the renal tissue of patients with lupus nephritis. This cause-effect relationship investigation of the damaged tissue is direct and more reliable than the conventional correlation analysis [19–20]. Gene–environment interactions add more complexity in explaining the etiology of autoimmune diseases. A recent study [21], using new computational methods, demonstrated that transcription factors (TFs) occupy multiple loci associated with individual complex genetic disorders. In particular, they showed that nearly half of systemic lupus erythematosus risk loci are occupied by the Epstein–Barr virus EBNA2 protein and many co-cluster human TFs, showing gene–environment interactions.

The following is a list of the other factors involved in SEL pathogenesis. Environmental factors include viruses, which stimulate specific cells in the immune system. Antibodies to the molecular mimicry molecules may contribute to the development of autoimmunity. In addition, trypanosomiasis or mycobacterial infections may induce anti-DNA antibodies or even lupus-like symptoms, and lupus flares may follow bacterial infections. UV light may stimulate keratinocytes to express more snRNPs and to secrete more IL-1, IL-3, IL-6, GM-CSF and TNF-alpha, thereby stimulating B cells to make more antibodies. Regarding hormonal factors, evidence of the immunoregulatory function of estradiol, testosterone, progesterone, dehydro-epiandrosterone, and pituitary hormones, including prolactin, has supported the hypothesis that they modulate the incidence and severity of SLE.

The use of estrogen-containing ocp is associated with a 50 percent increase in risk of developing SLE while either early onset of menarche (age  $\leq 10$  years) or administration of estrogen to postmenopausal women doubles their risk. Breastfeeding may decrease risk of developing SLE. Nulliparous women are at higher risk of SLE than are women who have given birth at least once. Treatment of women with clinically stable SLE with ocp for one year does not increase disease flares. However, treatment of postmenopausal women with hormone replacement may increase flares, although evidence is mixed. Hyperprolactinemia and hyperprogesteronemia are associated with flares of SLE.

SLE is primarily a disease characterized by abnormalities in immune regulation. There are numerous immune defects in patients with SLE. However, the etiology of these abnormalities remains unclear. We do not know which defects are primary, and which are secondarily induced. These abnormalities are thought to be secondary to a loss of self-tolerance. Thus, affected patients consequently develop an autoimmune response. B cells/plasma cells that make autoantibodies more persistently activated and driven to maturation by the B cell activating factor (BAFF, also known as B lymphocyte stimulator, blys) and by persistently activated T helper cells making B-supporting cytokines such as IL-6 and IL-10. Blys is essential for maturation and survival of post-bone marrow transitional and immature B cells into autoantibodysecreting plasmablasts and memory B cells. Antibody-Antigen complexes, particularly those containing DNA or RNA/proteins, activate the innate immune system via TLR9 or TLR7, respectively. Thus, dendritic cells are activated and release type 1 interferons and TNF-alpha, T cells release IFN-gamma, IL6, IL10, while natural killer (NK) and T cells fail to release adequate quantities of transforming growth factor (TGF)-beta. These cytokine patterns favor continued autoantibody formation. Phagocytosis and clearing of immune complexes, of apoptotic cells, and of necrotic cell-derived material are defective in SLE, allowing persistence of antigen and immune complexes. Immune complex may be present for years before the first symptom of disease appears. Blys production is promoted by increased TLR activation and increased type 1 and 2 interferon production. In turn, blys promotes increased TLR activation. This contributes to sustained autoantibody production. Blys increases survival of B2 cells after their transitional T1 phase, which bypass several deleting and energizing tolerance mechanisms [8].

A search of GeneCards database (http://www.genecards.org/) shows there are 1474 proteins associated with SLE. Comparing this list with a list of all the proteins in the 291 Gibbs-homology networks captured, we find an overlap of 12 proteins, namely: UBC, RPS6, RPS18, RPS10, RPLP2, RPL7, RPL6, RPL5, ISG15, IFS16, HSPA8, HSP90AA1. Interestingly and possibly importantly, these are all ribosomal proteins or proteins related to ribosomal proteins. Various reports have shown, that Anti-P-ribosomal are highly specific for SLE with a prevalence up to ~ 50%, while they are usually absent in other autoimmune disorders [22–23]. Anti-P-R target P0, P1 and P2 proteins are located on the eukaryotic ribosomal subunit and can penetrate cells and induce apoptotic changes, followed by an inhibition of specific cytokine [24–25]. It is worth noting that efficient ribosome biogenesis consumes over 60% of cellular energy supply in the form of ATP molecules and thus is strongly related to the energy status of the cell.

The "causal" factors result in T-, B-cell and dendritic cell dysfunctions of various types [26]. One of the important contributors to T cell dysfunction is the mitochondrial hyperpolarization, which results in ATP depletion, oxidative stress, and Ca<sup>+2</sup> and actin cytoskeleton depletion [27]. The T cells eventually rupture releasing pro-inflammatory nuclear materials [27]. This causes an imbalance in the T/B cell ratio, which promotes antibody production in the peripheral blood. This whole process cascades producing excessive necrotic debris. Dendritic cells sense the excess necrotic debris and soon inflammation is out of control. A typical therapy is to target B cells to allow normal processes to remove the cellular debris [5]. A more detailed description of this pathway is given by [2, 27–29]. Here, we focus on using molecular thermodynamics methodology in conjunction with systems biology to identify key proteins in this SLE inflammatory process. This may prove to be valuable for designing new treatments on a personalized basis for this pathology and other types of diseases as has already been attempted for cancer [30–32].

Finally, it should be mentioned that it would be very useful to correlate any thermodynamic measures that will be introduced in this paper with disease severity scores. Unfortunately, this is still very difficult. Measuring lupus disease activity accurately remains a challenging and demanding task. Many researchers have attempted to define what disease activity means and

how it should be measured, and several instruments were devised for a standardized assessment of disease activity and outcome domains in clinical research (i.e. BILAG-2004, SLAM-R, SLEDAI and SLEDAI-2 K, among others). Pitfalls involving lupus disease activity measures have had a significant impact on the interpretation of study outcomes. Many of the clinical trials were either overpowered or underpowered and had complex and multiple outcome measures. Devised in the 1990s, these metrics, however, were never intended for use in clinical trials [33]. The data set we used in the present study appears to contain patient data in an acute rheumatoid arthritis setting.

#### **Theoretical background**

The theoretical underpinnings for the thermodynamic approach to understand the molecular biology of human diseases were developed over a several-year period and involved different examples including several types of cancer [31-32,34-38]. Here, we give a brief summary of this body of work. The transcriptome and other -omic (e.g., proteomic, genomic, etc.) measures can be viewed as representing the energetic state of a cell. By the use of the word "energetic" we mean from a thermodynamics perspective. A living system is out of thermodynamic equilibrium simply because of a constant need for metabolic energy production. It uses nutrients such as glucose and transforms them into ATP as the universal biological energy currency required for structure formation and biological function. One of the energetic demands of every cell is the production of specific proteins, which are used for numerous structural and functional need of a cell. Protein expression levels, therefore, represent a measure of the living cell's non-equilibrium energy level. Moreover, proteins interact with other proteins generating very complex protein-protein interaction networks whose architecture is cell specific. There is a chemical potential between interacting molecules in a cell, and the chemical potential of all the proteins that interact with each other can be imagined to form a rugged landscape, not dissimilar to Waddington's epigenetic landscape [39-40]. The above formulates our conceptual framework for the foregoing analysis.

The method we propose uses mRNA transcriptome data or RNA-seq data as a surrogate for protein concentration. This assumption is largely valid. In fact, refs. [41–42] have shown an 83% correlation between mass spectrometry proteomic information and transcriptomic information for multiple tissue types. Further, ref. [43] found a Spearman correlation of 0.8 in comparing RNAseq and mRNA transcriptome from TCGA human cancer data (https://cancergenome.nih.gov/).

Given the set of transcriptome data, a representative of protein concentration, we overlay that on the human protein-protein interaction (PPI) network from BioGrid (<u>https://</u>thebiogrid.org/). This means we assign to each protein on the network, the transcriptome value (or RNAseq value) after rescaling. From that we then compute the Gibbs free energy of each PPI using the standard statistical thermodynamic relationship:

$$G_i = c_i \ln \frac{c_i}{\sum_j c_j} \tag{1}$$

where  $c_i$  is the "concentration" of the protein *i*, normalized, or rescaled, to be between 0 and 1 corresponding to minimum and maximum values, respectively. The sum in the denominator is taken over all protein neighbors of *i*, and including *i*. Therefore, the denominator can be considered akin to degree-entropy as pointed out elsewhere [31–32,34–38]. Carrying out this mathematical operation essentially transforms the "concentration" value assigned to each protein to a Gibbs free energy contribution. Thus, we replace the scalar value of transcriptome to a scalar function—the Gibbs free energy.

Due to the presence of a logarithm function of a fractional number, the Gibbs free energy is a negative number, so associated with each protein on the network is a negative free energy well (local energy minimum), which corresponds to a local stability area with respect to small changes in protein concentrations. This results in a rugged free energy landscape represented schematically in Fig 1. If we use what is called a topological filtration on this landscape, we essentially move a filtration plane up from the deepest energy well. As the filtration plane is moved up, larger-and-larger energetic subnetworks are captured. For convenience, we stop the filtration at energy threshold 32 –meaning 32 nodes in the energetic subnetwork are retained. We call these subnetworks Gibbs-homology networks.

We now compute the Betti centrality, which is a topological measure, of the 32-node energetic networks as described in detail earlier in ref. [36]. The main concept is easily described as follows. In networks such as PPI networks, there are holes, or rings, of various sizes. In these energetic pathways within PPI networks, the proteins form interaction rings. In densely connected, but not fully connected, networks the rings, or holes, may consist of triangles and larger rings of interaction. To find the Betti centrality we ask ourselves the following questions: which protein when removed from the network will change the overall total number of rings the most? The total number of rings is called the Betti centrality is given by the simple formula:

$$B(v_i) = B(G) - B(G - \{v_i\})$$
(2)

Hence, the difference from the total Betti number B(G) and the Betti number of the network after removing node *i*, gives the Betti centrality for node *i*. We then compute this property for all nodes in the threshold-32 energetic network. Often there will be two or more proteins in the network that have equivalent Betti centrality making them equally important to the network. We discuss this equivalence and the Betti centrality with respect to the patient data later in this manuscript.

#### Data source and methods

We use the Betti centrality measure, described above, on the Gibbs-homology network. The algorithm used for the calculation of Gibbs energy, Gibbs-homology and Betti centrality has been briefly described above and in detail elsewhere [34–38]. The dataset for this SLE study is from ref. [44], a study of lymphotoxin-Light pathway regulation by treatment of SLE and rheumatoid arthritis (RA) patients with baminercept [44]. The dataset is publicly available at GEO (https://www.ncbi.nlm.nih.gov/geo/) with the accession number GSE45219. This dataset





comprises data on 291 SLE patients and it is the largest of the datasets on SLE in the available literature. We believe it is highly representative. The dataset for the SLE study is available from GEO as accession number GSE45219. As pointed out in [23] the data are from whole blood samples not peripheral blood.

While the analysis we perform concerns protein, the data are derived from transciptional analysis of RNA. However, mRNA codes for proteins and the quantitative correlation between the two measures is very significant, see references [41-43]. We did not perform extensive statistical methodology in the paper, rather we focused on the markers of SLE [31-32,34-38]. In fact, the importance of specific proteins that our approach reveals is based on the methodology involving the Gibbs energy and topology of the protein subnetworks and not on statistics. This can be applied to individual patients in the framework of personalized therapy and that is a strength of this approach.

#### **Results and discussion**

As described above, we use the Gibbs-homology pathways, or small energetic networks, to find which protein is contributing most to the energetic pathway complexity. We do this by calculating the Betti number centrality. For some patients there will be one or more equivalent Betti number centralities. A Pareto chart of these "centrality proteins" for the 291 SLE patients in the GSE45291 dataset is shown in Fig 2.

The first thing to notice in this Pareto chart is that the number of proteins totals beyond 291 because of equivalent centralities. The second feature is that UBC, HUWE1 and RPS11 are present as centrality proteins in 100 or more patients, which is a very substantial number. Notice also the vertical axis is on a logarithmic scale, hence the actual differences are much larger than they appear.

A search of GeneCards database (http://www.genecards.org/) shows there are 1474 proteins associated with SLE. Comparing this list with a list of all the proteins in the 291 Gibbs-homology networks captured, we find an overlap of 12 proteins, namely: UBC, RPS6, RPS18, RPS10, RPLP2, RPL7, RPL6, RPL5, ISG15, IFS16, HSPA8, HSP90AA1. Interestingly and possibly importantly, these are all ribosomal proteins or proteins related to ribosomal proteins. It is



Fig 2. Pareto chart of highest Betti centralities for the Gibbs-homology energetic networks at threshold 32. In many networks there were one or more equivalent Betti centralities, so the total number of proteins is greater than 291, which is the number of SLE patients in the dataset. Note, the vertical axis is on a log scale.

worth noting that efficient ribosome biogenesis consumes over 60% of cellular energy supply in the form of ATP molecules and thus is strongly related to the energy status of the cell. This particular feature causes the nucleolar process to be highly sensitive to nutrient deprivation as demonstrated in recent studies on the target of rapamycin (TOR) signaling pathway, which plays a central role in linking the cellular nutrient status to ribosomal biogenesis [45]. In the following paragraphs we discuss each of these proteins and their role in SLE.

UBC (ubiquitin C), as to be expected, has a high Betti centrality in this population of patients, because it has high entropy. Entropy alone does not dictate high Betti centrality, but high entropy does play a role in the Gibbs energy calculation. As stated above, degree-entropy is essentially the denominator in Eq.(1). The degree-entropy for UBC in the Human Biogrid protein-protein interaction network version 3.4.139 (https://thebiogrid.org/) is 1432. This means it has 1432 protein neighbors with which it interacts. Here, entropy is specifically defined as the degree-entropy or the number of interactions. Similarly, HUWE1 has 455 neighbors and RPS11 (ribosomal protein S11) has 198 neighbors. It should also be noted that UBC, HUWE1 and RPS11 are all in the ubiquitination pathway. Fig 3, shows an example of one of the networks at energy threshold 32, in which UBC is the highest Betti centrality node. In this graph the proteins in the outer ring (HUWE1, RPL10, RPS20, HNRNPU, RPS11, NPM1, RPS3, RPL5, RPS8, HNRNPK) are all neighbors. As clearly seen in this graph, UBC does not have the highest degree entropy (but RPS8 does). Nonetheless, it has the highest Betti centrality for this patient.

RPS6 (ribosomal protein S6) was found in the GeneCards SLE list, and also found as having a high Betti centrality for 11 SLE patients out of the 291 total (Fig 3). However, its importance to SLE is indicated by the fact that it was found in 277 patients' Gibbs-homology pathways at energy threshold 32. So it does not necessarily have the highest Betti centrality but it is in the "energy neighborhood" at threshold 32 for 95% of the patients. RPS6 can be phosphorylated



**Fig 3. PPI network for patient 107, which is an example of UBC having the highest Betti centrality.** The nodes in the outer ring have direct connections to UBC. The nodes in the inner ring are secondary. RPS8 has the highest number of connections.

and is associated with the functioning of mTOR in T-cell development [46]. It also plays an important role in treatment of SLE with N-acetylcysteine, a target of which is mTOR [47]. Fig 4 shows a situation where RPS6 has a high Betti centrality and in fact has three other equivalent Betti centralities (RPS20, RPSA, RPL23A). Within the nodes in the circle, CUL3 has the highest number of connections. The peripheral nodes, nodes of secondary energy importance are: HUWE1 –associated with the ubiquitination pathway; HNRNPK—known as heterogeneous nuclear ribonucleoprotein K and according to KEGG (http://www.genome.jp/kegg/pathway. html) is involved in Herpes simplex infection and viral carcinogenesis; ACTB-known as actin beta is associated with Rap1 signaling pathway and Hippo signaling pathway, platelet activation, and according to OMIM (https://www.omim.org/) is associated with Baraitser-Winter syndrome; IQGAP1 -IQ motif containing GTPase activating protein—is associated with adherens junction and regulation of cytoskeleton, and is found 34 times at energy threshold 32 in our SLE patient dataset; IFI16 -interferon gamma inducible protein number 16 found 9 times at energy threshold 32. IFI16 is often up-regulated in SLE patients and it is suggested that it plays a key role in T-cell development. A low level of mRNA IFI16 expression has been found in naïve CD8<sup>+</sup> T-cells. However, CD8<sup>+</sup> mature cells express a high level of IFI16 mRNA [48]. It has also been reported that IFN-induced expression may depend on race, and the variation in coding regions of the polymorphs of the gene could account for differential regulation among individuals [48]. This implies IFI16 may be a good target for treatment by inhibition.

Many of the proteins and high-Gibbs energy nodes discussed above are commonly found in our analysis. Consequently, we will not repeat these individual protein commentaries.

RPS10 (ribosomal protein S10) has a high Betti centrality in 15 SLE patients from our database and at Gibbs energy threshold 32 is found in 277 of the 291 patients (Fig 4). RPS10 is often dimethylated in SLE patients [49].

RPS18 (ribosomal protein S18) was found to have high Betti centrality in the Gibbs homology networks for 6 SLE patients; and at energy threshold 32, it was found in 68 networks in the 291 SLE patients (Fig 5). Incidentally, RPS6 and RPS18 are both overexpressed in cancer [50]. In this connection it is worth mentioning that ribosomal proteins are known to control the







Fig 5. Patient number 30, RPS10 and RPS18 are equivalent high Betti centrality. The nodes in the ring are nearest neighbors to these two.

https://doi.org/10.1371/journal.pone.0226883.g005

expression and activity of key tumor suppressors including p53 [51] and a mutated in multiple cancer predisposition disorders, which are known as ribosomopathies [52]. A Gibbs-homology network graph showing both RPS10 and RPS18 as equivalent high Betti centralities is shown in Fig 5. These two nodes have as nearest neighbors all the nodes in the ring. The secondary nodes are: ACTB, UBC, HNRNPA1, EEF1A1, and HSP90AA1. HNRNPA1 has a high Betti centrality 7 times in our database and is present in 278 Gibbs homology networks at threshold 32. According to KEGG it, like HNRNPK, is involved in splicesome. Further, according to OMIM (omim.org) it contributes to Paget disease and amyotrophic lateral sclerosis. EEF1A1 (elongation factor 1 alpha promoter) is very commonly associated with prostate cancer [53], among many other cancers. It is also associated with oncogenesis, apoptosis and viral infections [54]. Of course it is not surprising to see HSP90AA1 in the energetic pathway. Heat shock proteins are often over-expressed in stress situations ranging from simple lesions to cancer. Its role in lupus is discussed in ref. [55].

In Fig 6 we see three, high Betti centrality nodes: RPL7, RPL7A and RPS19. All the proteins in the ring are neighbors of these three. RPL7 (ribosomal protein L7) was found to have a high Betti centrality in 10 patients and found to have high Gibbs energy 252 times in the database of 291 SLE patients. It is well- known that in SLE patients an autoimmune response to RPL7 is related to T cell activity [56]. RPL7A has a high Betti centrality only once out of 291 SLE patients. Furthermore, it has a high Gibbs energy (threshold 32) in 3 patients. RPS19 has a high Betti centrality 3 times and a high Gibbs energy 37 times. Very little can be found in the literature that describes RPS19 and SLE. According to OMIM it is associated with Diamond-Blackfan anemia. The two secondary energy nodes in Fig 6 are RPS13 and RPS3A. Interestingly, neither has a high Gibbs energy for 30 patients and RPS3A has a high Gibbs energy for 278 patients. These both are clearly important nodes for deeper investigation either as biomarkers or protein targets for inhibition.



**Fig 6. Patient number 124, all nodes in the circle have equivalent high Betti centrality.** RPS19, RPL7 and RPL7A have the highest Betti centrality in this graph.

https://doi.org/10.1371/journal.pone.0226883.g006

In Fig 7,RPL6 (ribosomal protein L6), and all those in the inner circle were found to have a high equivalent Betti centrality. RPL6 was found to have a high Betti centrality in 11 patients and to have a high Gibbs energy 250 times in the dataset of 291 SLE patients. It was also found in GeneCards as being important to SLE. The graph contains interferon-induc-ible family gene number 16 (IFI16). It is associated with SLE [48]. IFI16 was found in Gene-Cards as being important to SLE. We found it four times in our database of 291 patients. In patient #265 (see Fig 8) it is a neighbor of two high Betti centrality nodes, APP (highlighted) and RPL10 (not highlighted). It is also a neighbor of HUWE1 and HNRNPU. APP (Amyloid Precursor Protein) has a high Betti centrality. According to GeneCards it is not related to SLE but is associated with Alzheimer's disease. APP is found to have a high Gibbs energy 25 times in the database of 291 patients. Lastly, we point out that IQGAP1 is also in the energetic pathway for this patient.

RPL5 (ribosomal protein L5) was found to have a high Betti centrality in 10 patients and to have a high Gibbs energy in 21 times in the dataset of 291 SLE patients. Fig 8, patient number 49 is one of the patients with RPL5 as high Betti centrality. Very little research has been reported on RPL5, RPL6 and RPL7, with respect to SLE. Given their prevalence in our dataset (218, 250, 252 respectively), at energy threshold 32, it is reasonable to assume that these would be good proteins for further research into SLE treatment and mechanisms.

Three other proteins that should be discussed because of their frequency in our Gibbs energy analysis or because of their importance in SLE are: HSPA8, HSP90AA1 and ISG15. The two heat shock proteins HSPA8, HSP90AA1 are associated with many inflammatory diseases including rheumatoid arthritis, most cancers and SLE [32]. HSP90AA1 was found to have a high Betti centrality 13 times in the dataset of 291 SLE patients and found to have a high Gibbs energy (threshold 32) 278 times—almost every patient.

ISG15 interferon-stimulated gene 15, was found in GeneCards as being associated with SLE. More specifically, ref. [57] found it to be more highly expressed relative to healthy



Fig 7. Patient number 265, RPL6, among others in the inner circle are equivalent targets.

https://doi.org/10.1371/journal.pone.0226883.g007





controls (p = 0.032) in patients with SLE who had lymphocytopenia prior to treatment. ISG15 is believed to up-regulate macrophage migration inhibitory factor [58]. Further, it has 187 protein neighbors in the version of BioGrid PPI we used. It was twice highly-expressed in our dataset of 291 SLE patients, and was found to be highly expressed in four of the 492 rheumatoid arthritis patients.

Finally, we list ribosomal proteins and their functions that are all associated with immune signaling response and were found to have high Betti centrality: RPL13A, GAIT complex formation; RPS3, activation of NFkB; RPSA, target of tuberculosis drug PZA; RPS19, inhibition of MIF, ERK and NFkB, also interacts with hantavirus; RPS6, stabilizes LANA; RPS25, promotes virus production [35]. RPL13A is a negative regulator of inflammatory proteins, thus it likely plays an important role in SLE. Additionally, there are also metabolomics aspects of SLE that have been outside the scope of this study but should be the focus of a future investigation into personalized therapies for this pathology [59].

#### Conclusions

This paper reports the results of a computational study analyzing protein-protein interaction networks involved in Lupus, which is a unique auto-immune disease. SLE is a complex disease with many manifestations. Some patients exhibiting one set of indications and another patient exhibiting a different set. With regards to molecular network indications an interesting cross-species study was done with Murine and Human Lupus Nephritis [60]. The authors discovered 20 commonly shared network nodes reflecting pathologic process and cite CDR1, CD14, CCL5, CTSS, CXCL10, STAT1, CXCR5, IRF7, HCK, LYN, CFB, IFI11, GPNMB, ITGAM. Several genes in this list are known to be associated with immune function. In particular, ITAGM is well known to be associated with SLE. Another study [61], investigated transcription network modules in search for new biomarkers for SLE. They discovered new biomarkers indicative of disease progression. This paper was similar to [62], in as much as, it also discusses new biomarkers signatures to SLE. In another study [63], of significance, the investigators concluded that BAFF and interferon gens expression was high but gene signatures of these did not change with disease activity. In general, they found, that plasma cell gene expression corresponded to typical SLE serological markers.

In the present paper we analyzed patient-specific data publicly available through the GEO database. We used a combination of systems biology (protein-protein interaction network analysis via Betti number calculations) and statistical thermodynamics (via Gibbs homology with energy threshold filtering) approaches to obtain information, which has statistical significance. We found close to 300 proteins, which play substantial roles in the patient population of approximately the same size. However, on closer inspection, most commonly implicated proteins with major roles in the PPI networks belong to only a few special classes. The most important class consists of ribosomal proteins and ribosomal-related proteins. Next, ubiquitin and proteins belonging to the ubiquitination pathways have shown to Fig prominently in this dataset. Finally, heat-shock proteins have been found to be importantly involved in these pathologies. Some of the identified proteins should be considered for therapeutic inhibition as they clearly appear to be biomarkers for Lupus. It appears that that SLE is related to ribosomal stress and it is likely that epigenetic factors such as nutritional status and exposure to sunlight may be associated aspects in this response. While speculative, the fact that the high incidence of SLE among black African-American females may point to the involvement of mitochondrial ribosomes. While this paper has not proposed specific therapeutic approaches to SLE, the discussed identification of the potential targets for inhibitions is, in out opinion, a major step in this direction.

#### Acknowledgments

EAR and GLK acknowledge partial funding from CSTS-Healthcare, Toronto, Canada. JAT acknowledges funding support from NSERC (Canada).

#### **Author Contributions**

**Conceptualization:** Elizabeth J. Brant, Edward A. Rietman, Giannoula Lakka Klement, Jack A. Tuszynski.

Formal analysis: Edward A. Rietman, Marco Cavaglia.

Funding acquisition: Jack A. Tuszynski.

Methodology: Edward A. Rietman, Jack A. Tuszynski.

Project administration: Edward A. Rietman.

Supervision: Edward A. Rietman.

Writing - original draft: Elizabeth J. Brant, Edward A. Rietman.

Writing – review & editing: Edward A. Rietman, Giannoula Lakka Klement, Marco Cavaglia, Jack A. Tuszynski.

#### References

- Tsokos GC. Systemic Lupus Erythematosus. New Engl. J. Med. 2011; 365(22); 2110–2121. https:// doi.org/10.1056/NEJMra1100359 PMID: 22129255
- 2. Guerra SG, Vyse TJ, Cunninghame GDS. The genetics of lupus: a functional perspective. Arthr. Res. & Therapy. 2012; 14:211.
- Nobee A, Vailant AJ, Akpaka PE, Poon-king P. Systemic Lupus Erythematosus (SLE): A 360 Degree Review. Am. J. Clin. Med. Res. 2015; 3(4); 60–63.
- Budarf ML, Goyette P, Boucher G, Lian J, Graham RR, Claudio JO, et al. A target association study in systemic lupus erythematosus identifies multiple susceptibility alleles. Genes and Imm. 2011; 12 (1); 51–58.
- Lisnevskaia L, Murphy G, Isenberg D. Systemic Lupus erythematosus. Lancet. 2014; 384; 1878–1888. https://doi.org/10.1016/S0140-6736(14)60128-8 PMID: 24881804
- Xiao G, Zuo X, Epigenetics in systemic lupus erythematosus (Review). Biomed. Rep. 2016; 4:135– 139. https://doi.org/10.3892/br.2015.556 PMID: 26893827
- Hui-Yuen JS, Zhu L, Wong LP, Jiang K, Chen Y, Liu T.et al. Chromatin landscapes and genetic risk in systemic lupus. *Arthritis Res. Ther.* 2016; 18(1); 281. https://doi.org/10.1186/s13075-016-1169-9 PMID: 27906046
- Mukhyapranaprabhu M, Reddy J. Systemic lupus recent advances in management (undated online report).
- Li LJ, Zhao W, Tao SS, Li J., Xu S.Z., Wang, et al. Comprehensive long non-coding RNA expression profiling reveals their potential roles in systemic lupus erythematosus", *Cell Immunol.* 2017; 319; 17– 27. https://doi.org/10.1016/j.cellimm.2017.06.004 PMID: 28622785
- Dai Y, Zhang L, Hu C, Zhang Y. Genome-wide analysis of histone H3 lysine 4 trimethylation by ChIPchip in peripheral blood mononuclear cells of systemic lupus erythematosus patients", *Clin. Exp. Rheumatol.* 2010; 28(2); 158. PMID: 20483040
- Zhang Z, Shi L, Dawany N, Kelsen J, Petri MA, Sullivan KE. H3K4 tri-methylation breadth at transcription start sites impacts the transcriptome of systemic lupus erythematosus, *Clin. Epigenetics.* 2016; 8 (1); 14.
- Tsai CY, Hsieh SC, Lu MC, Yu CL. Aberrant non-coding RNA expression profiles as biomarker/bio-signature in autoimmune and inflammatory rheumatic diseases. J Lab Precis Med. 2018; 3(51);1–13.
- Rai G, Rai R, Saeidian AH, Rai M. Microarray to deep sequencing: transcriptome and miRNA profiling to elucidate molecular pathways in systemic lupus erythematosus. *Immunol Res.* 2016; 64(1); 14–24. https://doi.org/10.1007/s12026-015-8672-y PMID: 26188428

- Lu MC, Yu CL, Chen HC, Yu HC, Huang H Bin, Lai NS. Aberrant T cell expression of Ca2+ influx-regulated miRNAs in patients with systemic lupus erythematosus promotes lupus pathogenesis. *Rheumatol*ogy. 2014; 54(2) 343–348. https://doi.org/10.1093/rheumatology/keu322 PMID: 25172935
- Lai NS, Koo M, Yu CL, Lu MC. Immunopathogenesis of systemic lupus erythematosus and rheumatoid arthritis: the role of aberrant expression of non-coding RNAs in T cells. *Clin Exp Immunol.* 2017; 19(1); 261.
- Stypińska B, Paradowska-Gorycka A. Cytokines and microRNAs as candidate biomarkers for systemic lupus erythematosus. Int J Mol Sci. 2015; 16(10); 24194–24218. https://doi.org/10.3390/ ijms161024194 PMID: 26473848
- Hsieh SC, Tsai CY, Yu CL. Potential serum and urine biomarkers in patients with lupus nephritis and unsolved problems. Open Access Rheumatol. Res. Rev. 2016; 8; 81.
- Cardenas-Gonzalez M, Srivastava A, Pavkovic M, Bijol V, Rennke HG, Stillman IE, et al. Identification, confirmation, and replication of novel urinary microrna biomarkers in lupus nephritis and diabetic nephropathy. *Clin. Chem.* 2017; 63(9); 1515–1526. https://doi.org/10.1373/clinchem.2017.274175 PMID: 28667184
- Tsai C-Y, Lu M-C, Yu C-L. Can urinary exosomal micro-RNA detection become a diagnostic and prognostic gold standard for patients with lupus nephritis and diabetic nephropathy? J. Lab. Precis. Med. 2017; 2(11).
- Fu L, Jin L, Yan L, Shi J, Wang H, Zhou B, et al. Comprehensive review of genetic association studies and meta-analysis on miRNA polymorphisms and rheumatoid arthritis and systemic lupus erythematosus susceptibility. *Hum Immunol.* 2016; 77(1); 1–6. <u>https://doi.org/10.1016/j.humimm.2014.09.002</u> PMID: 25218914
- Harley JB, Chen X, Pujato M, Miller D, Maddox A, Forney C, et al. Transcription factors operate across disease loci, with EBNA2 implicated in autoimmunity. *Nat Genet.* 2018; 50(5); 699. <u>https://doi.org/10. 1038/s41588-018-0102-3</u> PMID: 29662164
- Yaniv G, Twig G, Shor DBA, Furer A, Sherer Y, Mozes O, et al. A volcanic explosion of autoantibodies in systemic lupus erythematosus: A diversity of 180 different antibodies found in SLE patients. *Autoimmun. Rev.* 2015; 14(1); 75–79. https://doi.org/10.1016/j.autrev.2014.10.003 PMID: 25449682
- Zandman-Goddard G, Chapman JSY. Autoantibodies involved in neuropsychiatric SLE and antiphospholipid syndrome. *Semin. Arthritis Rheum.* 2007; 36(5); 297–315. <u>https://doi.org/10.1016/j.</u> semarthrit.2006.11.003 PMID: 17258299
- Reichlin M. Cellular dysfunction induced by penetration of auto antibodies into living cells: cellular damage and dysfunction mediated by antibodies to dsDNA and ribo- somal P proteins. J. Autoimmun. 1998; 11(5); 557–561. https://doi.org/10.1006/jaut.1998.0219 PMID: 9802943
- 25. Sun KH, Tang SJ, Chen CY, Lee TP, Feng CK, Yu CL, et al. Monoclonal ribosomal P autoantibody inhibits the expression and release of IL-12, TNF-α and iNOS in activated RAW macrophage cell line. J. Autoimmun. 2005; 24(2);135–143. https://doi.org/10.1016/j.jaut.2005.01.002 PMID: 15829406
- Grammatikos AP, and Tsokos GC. Immunodeficiency and autoimmunity: lessons from systemic erythematosus. Trends in Molecular Medicine. 2012; 18(2);101–108. <u>https://doi.org/10.1016/j.molmed.2011</u>. 10.005 PMID: 22177735
- Perl A, Gergely P Jr, Nagy G, Koncz A, Banki K, Mitochondrial hyperpolarization: a checkpoint of T-cell life, death and autoimmunity. Trends Immunol. 2004; 25(7); 360–367. <u>https://doi.org/10.1016/j.it.2004</u>. 05.001 PMID: 15207503
- Perl A. Pathogenic mechanisms in systemic lupus erythematosus. Autoimmunity. 2010; 43(1); 1–6. https://doi.org/10.3109/08916930903374741 PMID: 20014960
- 29. Perl A. Systems biology of lupus: Mapping the impact of genomic and environmental factors on gene expression signatures, cellular signaling, metabolic pathways, hormonal and cytokine imbalance, and selecting targets for treatment. Autoimmunity. 2010; 43(1); 32–47. <u>https://doi.org/10.3109/08916930903374774 PMID: 20001421</u>
- Rietman EA, Platig J, Tuszynski JA, Lakka Klement G. Thermodynamic measures of cancer: Gibbs free energy and entropy of protein-protein interactions. J. Biol. Phys. 2016; 42(3); 339–50. <u>https://doi.org/10.1007/s10867-016-9410-y</u> PMID: 27012959
- Carels N, Tilli T, Tuszynski JA. A computational strategy to select optimized protein targets for drug development toward the control of cancer diseases. PLoS one. 2015; 10(1): e115054.
- Breitkreutz D, Hlatky L, Rietman E, Tuszynski JA. Molecular signaling network complexity is correlated with cancer patient survivability. PNAS USA. 2012; 109; 9209–9212. <u>https://doi.org/10.1073/pnas.</u> 1201416109 PMID: 22615392
- **33.** Mikdashi J, Nived O. Measuring disease activity in adults with systemic lupus erythematosus: The challenges of administrative burden and responsiveness to patient concerns in clinical research. *Arthritis Res. Ther.* 2015; 17(1); 183.

- Rietman EA, Scott JG, Tuszynski JA, Klement GL, Personalized anticancer therapy selection using molecular landscape topology and thermodynamics. Oncotarget. 2016; 8(12); 18735–18745.
- Hinow PR, Rietman EA, Omar SI, Tuszynski JA. Algebraic and Topological Indices of Molecular Pathway Networks in Human Cancers. Math Biosci. Eng. 2015; 12; 1289–1302. <u>https://doi.org/10.3934/mbe.2015.12.1289</u> PMID: 26775864
- Benzekry S, Tuszynski JA, Rietman EA, Lakka Klement G. Design principles for cancer therapy guided by changes in complexity of protein-protein interaction networks. Biol Direct. 2015; 10 (1); 32.
- Rietman EA, Karp RL, Tuszynski JA. Review and application of group theory to molecular systems biology. Theor. Biol. Med. Model. 2011; 8 (1); 21
- Rietman EA, Platig J, Tuszynski JA, Klement GL. Thermodynamic measures of cancer: Gibbs free energy and entropy of protein–protein interactions. *J. biol.phys.*2016; 42(3); 339–350. https://doi.org/ 10.1007/s10867-016-9410-y PMID: 27012959
- **39.** Huang S. On the intrinsic inevitability of cancer: From foetal to fatal attraction. Sem. Cancer Biol. 2011; 21 (3); 183–199.
- Moris N, Pina C, Arias AM. Transition states and cell fate decisions in epigenetic landscapes. Nat. Rev. Genetics. 2016; 17(11); 693–703. https://doi.org/10.1038/nrg.2016.98 PMID: 27616569
- Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, et al. A draft map of the human proteome. Nature. 2014; 509 (7502); 575–58. https://doi.org/10.1038/nature13302 PMID: 24870542
- Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM, et al. Mass-spectrometrybased draft of the human proteome. Nature. 2014; 509 (7502); 582–7. <u>https://doi.org/10.1038/</u> nature13319 PMID: 24870543
- Geo Y, Sheng Q, Li J, Ye F, Samuels DC, Shyr Y. Large scale comparison of gene expression levels by microarrays and RNAseq using TCGA data. PLoS One. 2013; 8(8), e71462. <u>https://doi.org/10.1371/journal.pone.0071462</u> PMID: 23977046
- Bienkowska J, Allaire N, Thai A, Goyal J, Plavina T, Nirula A, et al. Lymphotoxin-LIGHT Pathway Regulates the Interferon Signature in Rheumatoid Arthritis. PLoS One. 2014; 9(11); e112545. <u>https://doi.org/10.1371/journal.pone.0112545 PMID: 25405351</u>
- Zhou X, Liao W-J, Liao J-M, Liao P, Lu H. Ribosomal proteins' functions beyond the ribosome. J. Mol. Cell Biol. 2015; 7(2); 92–104.
- Salmond RJ, Brownlie RJ, Mayuhas O, Zamoyska R, Mechanistic Target of Rapamycin Complex 1/S6 Kinase 1 signals Influence T Cell Activation Independently of Ribosomal Protein S6 Phosphorylation. J. Immunology. 2015; 195 (10); 4615–4622.
- Lai ZW, Hanczko R, Bonilla E, Caza TN, Clair B, Bartos A, et al. N-Acetylcysteine Reduces Disease Activity by Blocking Mammalian Target of Rapamycin in T Cells from Systemic Lupus Erythematosus Patients. Arthritis & Rheumatism. 2012; 64(9), 2937–2946.
- Choubey D, Panchanathan R. Interferon-inducible *Ifi*200-family genes in systemic lupus erythematosus. Immunol. Lett. 2008; 119(1–2); 32–41. <u>https://doi.org/10.1016/j.imlet.2008.06.001</u> PMID: 18598717
- 49. Ren J, Wang Y, Ling Y, Zhang Y, Bao S, Xu Z. Methylation of Ribosomal Protein S10 by Protein-arginine Methyltransferase 5 Regulates Ribosome Biogenesis. J. Biol. Chem. 2010; 285(17); 12695–705. https://doi.org/10.1074/jbc.M110.103911 PMID: 20159986
- Wang W, Nag S, Zhang X, Wang M-H, Wang H, Zhou J, et al. Ribosomal Proteins and Human Diseases: Pathogenesis, Molecular Mechanisms, and Therapeutic Implications. Med. Res. Rev. 2015; 35 (2); 225–285. https://doi.org/10.1002/med.21327 PMID: 25164622
- Fierabracci A, Pellegrino M. The Double Role of p53 in Cancer and Autoimmunity and Its Potential as Therapeutic Target. Int. J. Mol. Sci. 2016; 17 (12); 1975.
- Raiser DM, Narla A, Ebert BL. The emerging importance of ribosomal dysfunction in the pathogenesis of hematological disorders. Leukemia & Lymphoma. 2014; 55(3): 491–500.
- 53. Scaggiante B, Dapas B, Bonin S, Grassi M, Zennaro C, Farra R, et al. Dissecting the expression of EEF1A1/2 genes in human prostate cancer cells: the potential of EEF1A2 as a hallmark for prostate transformation and progression. British J. of Cancer. 2012; 106 (1); 166–173.
- 54. Abbas W, Kumar A, Herbein G. The eEF1A proteins: at the crossroads of oncogenesis, apoptosis, and viral infections. Frontiers Oncol. 2015; 5; 75.
- 55. Shukla HD, Pitha PM. Role of Hsp90 in Systemic Lupus Erythematosus and Its Clinical Relevance. Autoimmune Dis. 2012; 2012.
- Donauer J, Wochner M, Witte E, Peter H-H, Schlesier M, Krawinkel U. Autoreactive human T cell lines recognizing ribosomal protein L7. Int. Immunol.1999; 11(2); 125–132. <u>https://doi.org/10.1093/intimm/</u> 11.2.125 PMID: 10069410

- Yuan Y, Ma H, Ye Z, Jing W, Jiang Z. Interferon-stimulated gene 15 expression in systemic lupus erythematosus. Z. Rheumatol, 2017; 77 (3); 256–262.
- Feng X, Chen W, Xiao L, Gu F, Huang J, Tsao BP, et al. Artesunate inhibits type I interferon-induced production of macrophage migration inhibitory factor in patients with systemic lupus erythematosus. Lupus. 2017; 26; 62–72. https://doi.org/10.1177/0961203316651738 PMID: 27230555
- Wu Tianfu, Xie Chun, Han Jie, Ye Yujin, Weiel Jim, et al. Metabolic disturbances associated with systemic lupus erythematosus. *PLoS One* 2012; 7(6): e37210. https://doi.org/10.1371/journal.pone. 0037210 PMID: 22723834
- Berthier CC, Bethunaickan R, Gonzalez-Rivera T, Nair V, Ramanujam M, Zhang W, et al. 2012. Crossspecies transcriptional network analysis defines shared inflammatory responses in murine and human lupus nephritis. J. Immunol. 2012; 189(2); 988–1001. https://doi.org/10.4049/jimmunol.1103031 PMID: 22723521
- Chaussabel D, Quinn C, Shen J, Patel P, Glaser C, Baldwin N, et al. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. *Immunity*. 2008; 29(1); 150– 164. https://doi.org/10.1016/j.immuni.2008.05.012 PMID: 18631455
- Streicher K, Morehouse CA, Groves CJ, Rajan B, Pilataxi F. Lehmann KP, et al. The plasma cell signature in autoimmune disease. Arthritis & rheumatology. 2014; 66(1); 173–184.
- **63.** Petri M, Fu W, Ranger A, Allaire N, Cullen P, Magder LS et al. Association between changes in gene signatures expression and disease activity among patients with systemic lupus erythematosus. *BMC medical genomics.* 2019; 12(1); 4. https://doi.org/10.1186/s12920-018-0468-1 PMID: 30626389