



Gene Expression Analysis Reveals Novel Shared Gene Signatures and Candidate Molecular Mechanisms between Pemphigus and Systemic Lupus Erythematosus in CD4⁺ T Cells

Tanya Sezin^{1*}, Artem Vorobyev^{2†}, Christian D. Sadik¹, Detlef Zillikens^{1,2}, Yask Gupta^{2†} and Ralf J. Ludwig^{1,2†}

¹Department of Dermatology, University of Lübeck, Lübeck, Germany, ²Lübeck Institute of Experimental Dermatology (LIED), University of Lübeck, Lübeck, Germany

OPEN ACCESS

Edited by:

Herman Waldmann,
University of Oxford, United Kingdom

Reviewed by:

Huanfa Yi,
Jilin University, China
Anne Fletcher,
Monash University, Australia

*Correspondence:

Tanya Sezin
tanya.sezin@uksh.de

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Immunological Tolerance and
Regulation,
a section of the journal
Frontiers in Immunology

Received: 30 August 2017

Accepted: 22 December 2017

Published: 17 January 2018

Citation:

Sezin T, Vorobyev A, Sadik CD,
Zillikens D, Gupta Y and Ludwig RJ
(2018) Gene Expression Analysis
Reveals Novel Shared Gene
Signatures and Candidate
Molecular Mechanisms between
Pemphigus and Systemic Lupus
Erythematosus in CD4⁺ T Cells.
Front. Immunol. 8:1992.
doi: 10.3389/fimmu.2017.01992

Pemphigus and systemic lupus erythematosus (SLE) are severe potentially life-threatening autoimmune diseases. They are classified as B-cell-mediated autoimmune diseases, both depending on autoreactive CD4⁺ T lymphocytes to modulate the autoimmune B-cell response. Despite the reported association of pemphigus and SLE, the molecular mechanisms underlying their comorbidity remain unknown. Weighted gene co-expression network analysis (WGCNA) of publicly available microarray datasets of CD4⁺ T cells was performed, to identify shared gene expression signatures and putative overlapping biological molecular mechanisms between pemphigus and SLE. Using WGCNA, we identified 3,280 genes co-expressed genes and 14 co-expressed gene clusters, from which one was significantly upregulated for both diseases. The pathways associated with this module include type-1 interferon gamma and defense response to viruses. Network-based meta-analysis identified RSAD2 to be the most highly ranked hub gene. By associating the modular genes with genome-wide association studies (GWASs) for pemphigus and SLE, we characterized IRF8 and STAT1 as key regulatory genes. Collectively, in this *in silico* study, we identify novel candidate genetic markers and pathways in CD4⁺ T cells that are shared between pemphigus and SLE, which in turn may facilitate the identification of novel therapeutic targets in these diseases.

Keywords: autoimmunity, gene expression analysis, weighted gene co-expression analysis, pemphigus, systemic lupus erythematosus, CD4⁺ T cells

INTRODUCTION

Pemphigus is a rare autoimmune bullous dermatosis, clinically characterized by intraepidermal blistering of the skin and/or mucous membranes. Immunologically, pemphigus is characterized by autoantibodies directed against desmosomal and non-desmosomal adhesion molecules expressed in the skin and mucosa. Binding of the pathogenic autoantibodies in the skin leads to dissociation of adjacent keratinocytes and formation of blisters. Based on the clinical presentation and the specificity of the anti-desmoglein (Dsg) autoantibodies, pemphigus is classified into two main forms, pemphigus

vulgaris (PV), with autoantibodies targeting Dsg3, and in some cases also Dsg1, and pemphigus foliaceus (PF), with autoantibodies targeting Dsg1 (1). The association of pemphigus with connective tissue diseases such as systemic lupus erythematosus (SLE) has been previously noted on a case report/case series basis (2, 3). In line, pemphigus autoantibodies and antinuclear autoantibodies, one immunological hallmark of SLE (4), coexist in healthy blood donors (5). However, the molecular mechanism remains unknown. The co-occurrence of pemphigus and SLE can suggest a common network of multifunctional genes and pathways. Alternatively, it can be altogether serendipitous. Due to the complexity of such a system, weighted gene co-expression network analysis (WGCNA) can serve as a comprehensive tool for identifying gene clusters of correlating and connected shared genes (6, 7). This approach has been previously successfully applied in various biological contexts to identify regulatory genes and networks associated with multiple disease phenotypes (8–11).

Systemic lupus erythematosus and pemphigus are characterized by the production of autoantibodies and traditionally classified as B-cell-mediated autoimmune diseases. Compelling evidence has, however, shown that autoreactive helper-T lymphocytes are crucial in pathogenicity of both diseases by regulating B cells response and promoting autoantibodies production (12–15). Thus, studying gene expression networks within the CD4⁺ T-cell population is not only essential for understanding the underlying pathophysiology but also for identifying predictive biomarkers and establishment of novel therapeutic targets for these diseases.

Using publically available gene expression data from NCBI GEO database, we investigated gene co-expression networks of CD4⁺ T cells obtained from pemphigus (PV as well as PF) and SLE patients (16). Our analysis revealed 14 distinct modules

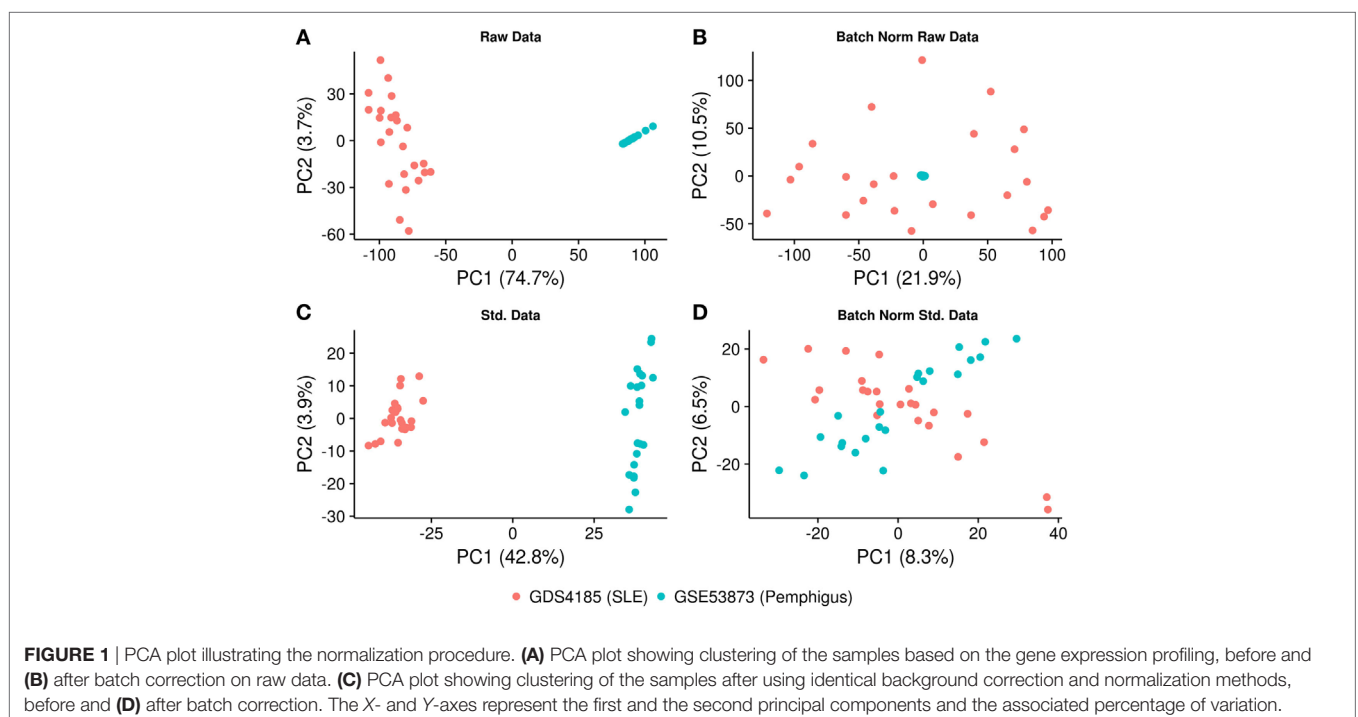
containing 3,280 co-expressed genes between the two diseases. Two out of 14 modules were found significantly upregulated: one in PF and SLE, and the other in PV. We further identified biological pathways such as “type I interferon signaling pathway” and “defense response to virus” using KEGG database, to be enriched in disease-associated modules. To the best of our knowledge, this is the first study applying a systems biology approach to identify shared molecular mechanisms between pemphigus and SLE.

MATERIALS AND METHODS

Data Collection

All the data for the analysis were collected by searching expression databases such as NCBI GEO and Array Express for CD4⁺ T cells for pemphigus and SLE (17, 18). The datasets from other tissues or cell type were discarded. Also, the datasets, which did not have raw data files, were discarded from the downstream analysis. Two datasets, one for pemphigus (GSE53873) and one for SLE (GDS4185), were included in this study. The covariate information available for the patients is summarized in Table S1 in Supplementary Material. Altogether 46 samples (4 PV, 15 PF, 13 SLE, and 14 healthy controls) were used in the analysis.

To avoid a potential bias that could be introduced by obtaining two separate microarray datasets, the deposited gene expression data were directly used for batch normalization. The expression profiles were log₂ transformed and batch normalization was done using “sva” and “combat” functions in SVA R package (19). The effect of normalization was investigated by principal component analysis using the R-based “prcomp” function. Since batch normalization still produced biased results (**Figure 1**), the raw files were preprocessed again and an additional normalization



step was performed. In detail, raw gene expression profiles were deduced from text files (Codelink array) using Codelink R package (20). Using the same package, first, the background was corrected with the “normexp” method and then normalized by the “cyclicloess” method. For Affymetrix data, raw gene expression for each sample was derived using R Affy package (21). The background correction was performed by “backgroundCorrect (method = ‘normexp’)” and cyclic normalization was performed on log₂ expression values using limma R package (22). All the probes from each of the microarray platforms were filtered out for significant low expression/variation ($P < 0.05$) using the “varianceBasedFilter” function from DCGI R package (23). The remaining probes were mapped to Ensembl gene identifiers and probes’ expression was collapsed to gene-level expression using “collapseRows” function with default parameters in WGCNA R package (24). Consequently, batch normalization and statistical analysis were performed on the overlapping genes between two platforms using “combat” and PCA analyses, respectively (25). The data were further investigated for the presence of confounding effects such as clinical form of the disease (generalized vs. localized) and treatment group (prednisone vs. untreated) for pemphigus dataset (GSE53873) using anosim function with 999 permutations in vegan R package (26).

Co-Expression Networks

Co-expression modules were generated using WGCNA R package. A signed weighted adjacency matrix of pair-wise connection strengths (bicor correlation) was constructed using the soft-threshold approach with a scale-independent topological power $\beta = 6$. For a gene, the connectivity was defined as the sum of all connection strengths with all other genes. Genes were aggregated into modules by hierarchical clustering and refined by the dynamic tree cut algorithm. Thereafter, module eigenvalues were calculated. The eigenvalue is the first principal component of the gene expression profile within a module, representing average module expression profile (27). The statistical significance ($P < 0.05$) of module eigenvalues among the groups was accessed by Kruskal–Wallis test. Modular hub gene candidates were identified by correlating the gene expression with its module eigenvalues (“chooseTopHubInEachModule” function in WGCNA). To generate the causal network within a module, the C3NET R package was used (28). The algorithm uses mutual information theory to construct gene networks from gene expression data. The final network was generated using “c3net” function with default setting. A gene–gene interaction was considered to be significant if $\alpha < 0.05$.

Functional Characterization of a Module

To investigate known gene–gene interactions, we used the INMEX web server (29). All genes within a specific module were queried, and a minimum network connecting all genes within this module was obtained. The hub gene candidates from this analysis were defined by their degree of interactions. Gene ontology terms, enriched KEGG pathways, and transcription factor binding sites for each module were obtained using David web server. Thereafter, all the mapped genes and reported genes to the disease-associated loci were selected from genome-wide association study (GWAS)

catalog. The selected genes and modular genes were connected to each other based on known gene–gene interactions (INMEX web server). Only the direct interactions between the modular genes and GWAS genes were considered. Gene–gene interactions were visualized using Cytoscape software and figures were generated using R programming language. Intermediate gene conversions and data formatting were done using Perl programming language (30).

RESULTS

Data Selection and Normalization

Microarray data were obtained for peripheral CD4⁺ T-cell samples from 19 pemphigus patients (4 PV; 15 PF), 13 SLE patients, and 14 healthy controls from NCBI GEO and EBI Array Express (GSE53873; GDS4185). Altogether, our dataset included 46 samples derived from Codelink and Affymetrix arrays. Only datasets comprising raw files were included in the downstream meta-analysis. Therefore, we excluded samples GSE4588 and GSM260948 from our analysis.

To implement the co-expression network analysis, we standardized and batch-normalized the datasets. We collected common probes across the two chip-arrays. The CodeLink Human Whole Genome Bioarray from GE Healthcare consisted of 54,359 probes, while the Affymetrix Human Genome U133A array consisted of 22,283 probes. We converted these probes to ensemble gene identifiers using ensemble biomaRt and found that 12,980 genes were common between the two platforms. Consequently, the datasets were merged based on the expression of common genes and “combat” and “sva” (SVA R package) functions were applied to remove the batch effect. Our results show that while the Affymetrix samples were distributed uniformly among the principal components, the data generated from the CodeLink array still clustered together (**Figures 1A,B**), suggesting that the dataset was not properly normalized and required further optimization. To further optimize the datasets, we used the “normexp” method for background correction and “cyclicloess” on log₂ transformed values. Additionally, each dataset was separately filtered for low expressing/varying probes, as well as multiple probes were collapsed for each gene. Briefly, 18,038 probes representing 12,980 genes were identified in the CodeLink dataset. These probes were filtered for low variation and collapsed to generate 5,646 gene expression profiles. Similarly, the Affymetrix gene chip consisted of 20,366 probes representing 12,980 genes. These probes were filtered and collapsed, resulting in 6,073 gene expression profiles. Overall, the overlap between the two datasets consisted of 3,280 gene expression profiles, which were further used in the downstream analysis. After applying the batch effect normalization “combat” algorithm, we observed that the samples were distributed among first principal component with only 8.3% variation explained by the first component (**Figures 1C,D**). We also analyzed confounding effects by stratifying the dataset for different covariates. We found no significant differences for covariate generalized vs. localized ($P = 0.402$) and prednisone treated vs. untreated ($P = 0.596$) for pemphigus samples. No covariate information was available for SLE samples (Figure S1 in Supplementary Material).

Detection of Co-Expression Modules Related to Pemphigus and SLE

Next, we set out to identify system-level similarity between pemphigus and SLE. Therefore, we applied WGCNA, aiming to identify gene modules that are co-expressed between pemphigus and SLE samples, and that are likely to be involved in common pathways. The major advantage of using such an approach is that it alleviates the multiple-testing problem that is inherent to microarray datasets. Using WGCNA, we identified 14 modules of co-expressed genes for 3,280 highly expressed and varying gene expression profiles, which are represented by different color codes (**Figure 2**; Figure S1, Data Sheet 1 in Supplementary Material). Two out of 14 modules showed differences between control and disease samples. The module “magenta” was significantly upregulated for both PF ($P = 0.005$) and SLE ($P = 0.016$) in comparison to healthy controls, and the module “salmon” was specifically upregulated only in PV ($P = 0.034$) (**Figure 2**).

Biological Pathways in the PF- and SLE-Associated Module “Magenta”

Module “magenta” consisted of 74 genes and, compared with controls, was significantly upregulated in PF and SLE. To investigate different known mechanisms associated with this module, we performed gene ontology analysis using DAVID database (31). We found that this module was, among others, enriched in biological processes such as “type I interferon signaling pathway” ($P_{\text{adj}} = 6.4E-11$), “defense response to virus” ($P_{\text{adj}} = 2.7E-10$), and “cytokine-mediated signaling pathway” ($P_{\text{adj}} = 1.3E-7$) (**Table 1**). This module was also enriched in KEGG pathways, including “measles” ($P_{\text{adj}} = 2.3E-4$), “influenza A” ($P_{\text{adj}} = 2.7E-4$), and “herpes simplex infection” ($P_{\text{adj}} = 1.3E-3$). On the basis of statistical module membership and eigengenes value, we identified s-adenosyl methionine

domain containing 2 (RSAD2) gene as the most highly ranked hub gene for this module. To identify subnetworks and statistical interactions within the modules we used the “c3net” algorithm. The “c3net” algorithm investigates the direct physical interaction for gene expression data, further providing putative mechanisms within a module and characterizing its key regulating genes (9). We found 2'-5'-oligoadenylate synthetase 1 (OAS1), MX dynamin-like GTPase 1 (MX1), interferon-induced protein with tetratricopeptide repeats 3 (IFIT3), and spermatogenesis-associated serine-rich 2 like (SPATS2L) genes as master regulator genes of the module ($\text{degree} \geq 5$) (**Figure 3**). Moreover, to further explore known gene–gene interactions among the genes in “magenta” module, we used the INMEX web server (32). We were specifically interested in examining “minimum interaction networks.” In this type of networks, a minimum number of genes are required to connect all the nodes to a given set of genes. Using this approach, we further derived additional regulators such as junction plakoglobin (JUP), B-cell CLL/lymphoma 2 (BCL2), ISG15 ubiquitin-like modifier (ISG15), STAT1, S-phase kinase-associated protein 2 (SKP2), and eukaryotic translation initiation factor 2 alpha kinase 2 (EIF2AK2) (Figure S2 in Supplementary Material).

Biological Pathways in the PV-Associated Module “Salmon”

Although the sample size for PV samples was small ($n = 4$), we identified a distinct module that, compared with controls, was significantly upregulated in PV, namely the “salmon” module ($P = 0.034$). The “salmon” module comprises 39 genes (**Table 1**) and was enriched in the following biological processes: “blood coagulation” ($P_{\text{adj}} = 1.4E-1$) and the KEGG pathway “platelet activation” ($P_{\text{adj}} = 1.8E-1$). Using statistical module eigengenes, we identified platelet glycoprotein IX (GP9) as a hub gene of this

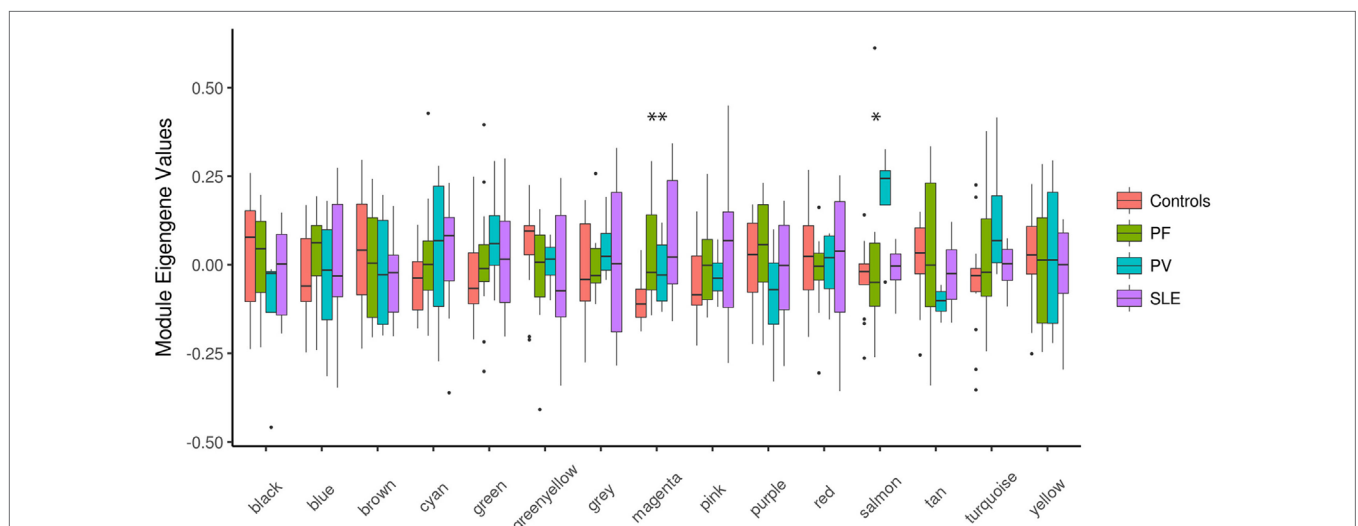
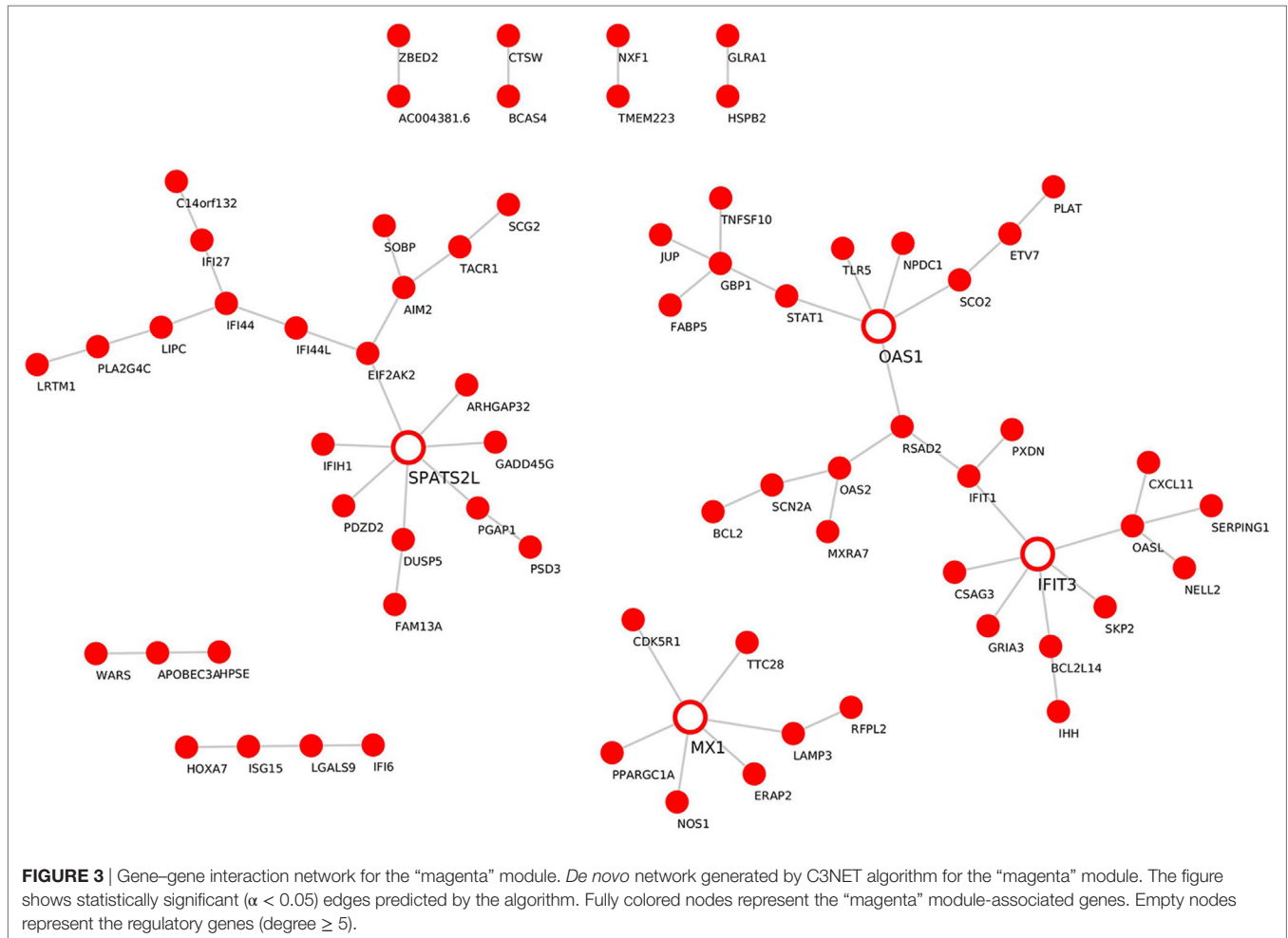


FIGURE 2 | Boxplots of eigengene values across modules. Boxplots depicting different identified modules on the X-axis and the corresponding module eigengene values for each group of samples on the Y-axis. The significance among the groups was calculated using Kruskal–Wallis test. * $P < 0.05$; ** $P < 0.01$. PF, pemphigus foliaceus; PV, pemphigus vulgaris; SLE, systemic lupus erythematosus.

TABLE 1 | Gene ontology and enriched KEGG pathways for “magenta” and “salmon” modules.

Module	Category	Term	P-value	Benjamini
Magenta	UP_KEYWORDS	Antiviral defense	1.18273E-16	1.84297E-14
	UP_KEYWORDS	Immunity	1.22704E-13	1.01824E-11
	GOTERM_BP_DIRECT	GO:0060337~type-I interferon signaling pathway	9.37804E-14	6.3981E-11
	UP_KEYWORDS	Innate immunity	3.82091E-12	2.11426E-10
	GOTERM_BP_DIRECT	GO:0051607~defense response to virus	7.83394E-13	2.6713E-10
	GOTERM_BP_DIRECT	GO:0045071~negative regulation of viral genome replication	1.21675E-10	2.76607E-08
	GOTERM_BP_DIRECT	GO:0009615~response to virus	2.90413E-10	4.95154E-08
	GOTERM_BP_DIRECT	GO:0019221~cytokine-mediated signaling pathway	9.178E-10	1.25188E-07
	KEGG_PATHWAY	hsa05162:Measles	4.89228E-06	0.00022502
	KEGG_PATHWAY	hsa05164:Influenza A	2.9062E-06	0.000267335
	KEGG_PATHWAY	hsa05168:Herpes simplex infection	4.20496E-05	0.001288717
	GOTERM_MF_DIRECT	GO:0003725~double-stranded RNA binding	6.2164E-05	0.009466294
	GOTERM_BP_DIRECT	GO:0060333~interferon-gamma-mediated signaling pathway	0.000216281	0.024286767
Salmon	GOTERM_BP_DIRECT	GO:0030041~actin filament polymerization	0.000889415	0.183621158
	GOTERM_BP_DIRECT	GO:0007596~blood coagulation	0.001317229	0.139518478
	KEGG_PATHWAY	hsa04611:Platelet activation	0.003518509	0.179124611



module. Additionally, using the “c3net” algorithm, we identified pro-platelet basic protein (PPBP), G protein subunit gamma 11 (GNG11), and thrombospondin 1 (THBS1) genes as key regulators of the “salmon” module (degree ≥ 4) (Figure 4). In addition,

while using the INMEX server we identified protein kinase cAMP-dependent type-II regulatory subunit beta (PRKAR2B), Src homology 2 domain-containing-transforming protein 3 (SHC3), tensin 1 (TNS1), PPBP, and GNG11 as regulatory genes

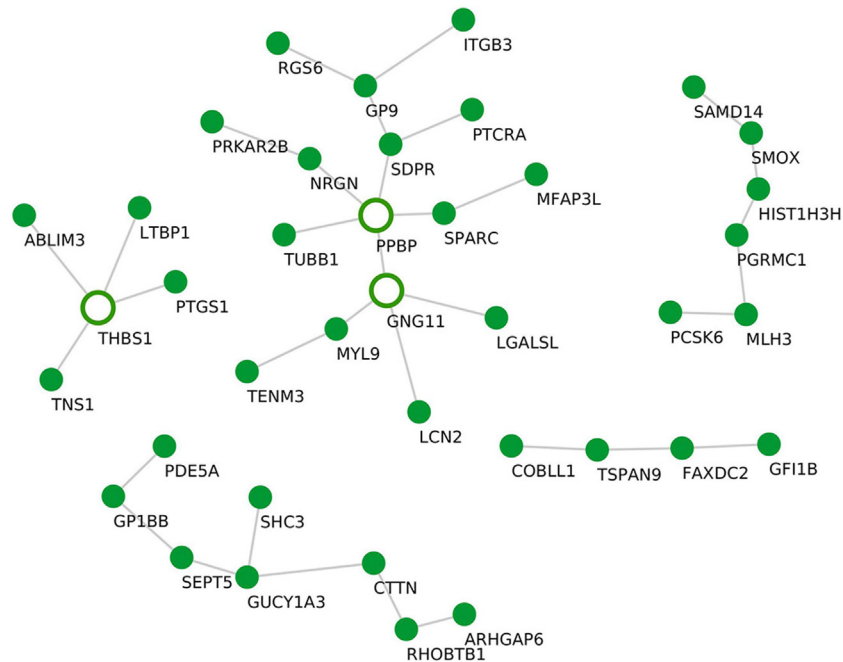


FIGURE 4 | Gene-gene interaction network for the “salmon” module. *De novo* network generated by C3NET algorithm for the “salmon” module. The figure shows statistically significant ($\alpha < 0.05$) edges predicted by the algorithm. Fully colored nodes represent the “salmon” module-associated genes. Empty nodes represent the regulatory genes (degree ≥ 4).

(Figure S3 in Supplementary Material). Interestingly, both PPBP and GNG11 genes coincided with the list of the aforementioned C3NET-derived key regulatory genes.

Cross-Linking SLE and Pemphigus GWA Studies with Clusters of Co-Expressed Genes in the “Magenta” and the “Salmon” Modules

While multiple GWA studies had been undertaken in a continuous effort to identify SLE susceptibility genes, only one GWA study was previously conducted in pemphigus, namely in PV (33, 34). In contrast to GWA studies that normally investigate the causal genes for a disease phenotype, gene expression profiles indicate the downstream effector phase. In the present work, we investigated direct interactions between previously reported susceptibility genes in SLE and pemphigus GWA studies and genes comprising the “magenta” and “salmon” modules, which were identified herein. We found the SLE-susceptible gene interferon regulatory factor 8 (IRF8) to have the largest number of direct interactions with “magenta” module-associated genes (Figure 5). The IRF8 gene interacted with genes encoding for interferon-induced protein with tetratricopeptide repeats 1 (IFIT1), interferon-induced guanylate-binding protein 1 (GBP1), 2'-5'-oligoadenylate synthetase 2 (OAS2), 2'-5'-oligoadenylate synthetase-like (OASL), and signal transducer and activator of transcription 1 (STAT1). Both IRF5 and STAT1 SLE GWAS genes directly interacted with IRF8 and with the other 4 “magenta” module-associated genes such as interferon induced with helicase

C domain 1 (IFIH1), IFIT1, GBP1, OASL, OAS2, and EIF2AK2 (Figure 5). Polymorphism in the gene ST18 has been previously found in a PV GWA study. However, we could not identify direct interactions between ST18 and genes associated with the “salmon” module. To further establish a putative association of ST18 to other genes in the “salmon” module, we performed the transcriptional factor binding sites enrichment analysis (39 “salmon” genes and the ST18 gene). We observed that 34 out of the 40 analyzed genes are regulated by the nuclear hormone peroxisome proliferator activated receptor γ (PPAR- γ ; P . adj = $8.3E-3$) and 25 out of 40 genes are regulated by growth factor independent 1 transcriptional repressor (GFI1; P .adj = $8.3E-3$).

DISCUSSION

The pathogenesis of most autoimmune disorders is still largely unknown. Environmental triggers in genetically susceptible individuals, as well as molecular mimicry mechanisms, may only partially account for this phenomenon (35). The co-occurrence of autoimmune diseases has been previously documented and aided in our understanding of autoimmunity (36).

Pemphigus and SLE are well-characterized autoimmune diseases that were previously reported to coexist in the same patient (37). Even though each of these two autoimmune diseases affects distinct organs and systems, the comorbidity of both diseases suggests an existence of fundamental common pathophysiological mechanisms. As we were interested in systems level similarity between the diseases rather than characterizing individual gene signatures, we used WGCNA to study pemphigus and SLE. Using this analytical approach, we identify modules across

patients may present severe systemic viral infections primarily associated with Epstein-Bar virus (EBV), cytomegalovirus, and herpes simplex virus (HSV). With respect to pemphigus, in 1974, Krain et al. first reported the association between HSV and PV (43); meanwhile, several additional case reports were published examining this association (44–46). A more recent study by Kurata and colleagues demonstrated high levels of HSV DNA in the saliva of PV patients at the earliest stage of the disease without a history of herpetic infection, thus suggesting the presence of cases of pemphigus induced by herpesviruses (47).

In our work, on the basis of statistical module membership and its eigengene value, we identified RSAD2 gene as the hub gene of the “magenta” module. Notably, by examining the expression levels of RSAD2 gene in our datasets we could demonstrate its significant upregulation in PF ($P = 0.005$) and SLE ($P = 0.007$) in comparison to healthy controls (Figure S4A in Supplementary Material). To confirm, the expression of the RSAD2 gene is encoding for interferon-inducible viperin protein, which inhibits viral replication and facilitates T-cell receptor-mediated GATA3 activation, and optimal Th2 cytokine production through modulation of NFKB1 and JUNB activities. As a result, viperin-deficient mice show impaired Th2 cell development (48). Interestingly, transcripts for RSAD2 were found to be upregulated in SLE CD3⁺ CD4⁺ cells, as well as SLE CD19⁺ B cells, and SLE CD33⁺ myeloid cells in comparison to similar cellular subsets isolated from healthy controls (49). Although it has been previously demonstrated that Th2 cells exert broad activity in blister formation in pemphigus, the association of RSAD2 with pemphigus is unknown. To examine the relevance of Th2 response in pemphigus and SLE, a set of 44 genes associated with Th2 differentiation were downloaded from the PathCards Pathway Unification Database from the Weizmann Institute of Science, and examined for their fold change expression in our disease datasets (PV, PF, and SLE) in comparison to healthy controls (Figure S4B in Supplementary Material). Our findings confirm that the fold change expression of Th2-associated genes was positively correlated between SLE and PF ($P = 0.01$, $\rho = 0.36$) and between SLE vs. PV ($P = 1.087E-05$, $\rho = 0.62$), suggesting that the Th2 response is skewed in a similar pattern between SLE and pemphigus. While investigating subnetworks within the “magenta” module (using the “c3net” algorithm), we identified OAS1, MX1, IFIT3, and SPATS2L genes as master regulators (Figure 3). Additional regulatory genes such as JUP, BCL2, ISG15, STAT1, SKP2, and EIF2AK2 were identified using known gene–gene interactions database (INMEX) (Figure S2 in Supplementary Material). Transcripts of 7 out of the 11 identified genes (i.e., RSAD2, OAS1, MX1, IFIT3, ISG15, STAT1, and EIF2AK2) were previously shown to be upregulated in SLE CD3⁺ CD4⁺ cells (49). Consistent with a previous study that examined possible related signaling pathways shared in the pathogenesis of several systemic autoimmune diseases (SAID) such as dermatomyositis, polymyositis, rheumatoid arthritis, and SLE, a subset of five viral-related differentially expressed genes (i.e., RSAD2, IFIT3, ISG15, STAT1, and EIF2AK) was detected in peripheral blood of SAID probands and their unaffected twins (50). Additionally, other genes that were identified in our study, including BCL2, OAS1, MX1, and SKP2 have been previously associated with various autoimmune diseases (51–54). Therefore,

our findings further suggest that these common IFN signature genes are shared across multiple autoimmune diseases including pemphigus and SLE.

Here, we identified a PV-specific associated module. The “salmon” module consisted of 39 genes and was enriched in genes involved in blood coagulation and platelet activation. Based on the eigengene value, the gene GP9 was identified as the hub gene of the “salmon” module. GP9 encodes a small-membrane glycoprotein that is part of the GPIb-V-IX complex that mediates platelet adhesion to blood vessels and promotes hemostasis. Thus, mutations in the GP9 protein lead to a coagulation disorder, also known as the Bernard–Soulier syndrome, characterized by thrombocytopenia. Of note, although this is a first report suggesting a role for GP9 in PV, a previous study by Hunziker et al. identified platelet-derived factors to enhance pemphigus acantholysis in skin organ cultures (55). Moreover, another study by Mizutani et al. found increased expression of the coagulation factor on keratinocytes, which shield blisters in PV (56). In line with this observation, using the “c3net” algorithm, we identified an additional list of platelet-associated genes i.e., PPBP, GNG11, and THSB1, as key regulators of the “salmon” module (Figure 4). Furthermore, by examining known gene–gene interactions, we could identify PPBP, GNG11, as well as another group of platelet-function-associated genes such as PRKAR2B, SHC3, and TNS1 (Figure S3 in Supplementary Material) as additional regulators of this module.

Further in our analysis, we associated the genes found in the “magenta” and “salmon” modules with known susceptibility markers of PV and SLE, which had been formerly identified by GWASs. GWASs are applied to identify genetic variants that are associated with a disease trait. However, the identification of loci harboring the susceptible genes does not fully reveal the molecular mechanisms that are at play to yield the observed phenotype. Therefore, linking these susceptibility genes with the module-associated genes may identify pathways that control the disease phenotype and provide potential therapeutic targets for intervention. By cross-linking susceptibility genes derived from SLE GWAS with clusters of co-expressed genes in “magenta” module, we found IRF8 to directly interact with the largest number of interferon-induced genes present in the “magenta” module including IFIT1, GBP1, OAS2, OASL, and STAT1 (Figure 5). Interestingly, STAT1 was identified both as an SLE susceptibility gene and as a key regulator gene of the “magenta” module. Therefore, based on our analysis, we predict IRF8 to have pharmacological relevance, as previously described (57). With regard to PV, we did not identify direct interactions between the known GWAS gene, ST18, and the 39 “salmon” module-associated genes. To circumvent this finding, we additionally performed a transcriptional factor binding sites enrichment analysis for the 40 genes. We found that the majority of the genes are regulated by the transcription factors PPAR- γ and GFI1 that have been previously described for their role in Th2 cell development (58, 59). Moreover, PPAR- γ has been suggested as a pharmacological target for PV (60).

Altogether, our work reveals conserved molecular mechanisms and pathways between pemphigus and SLE and identifies novel gene candidates that could be used as biomarkers or as potential targets for therapeutic intervention.

AUTHOR CONTRIBUTIONS

TS, AV, YG, and RL designed the study, interpreted the data, and wrote the manuscript. All authors contributed equally to this work. YG downloaded and analyzed the data. CS and DZ discussed the results and contributed to the writing of the manuscript.

ACKNOWLEDGMENTS

We thank Prof. SM Ibrahim (*Lübecker Institut für Experimentelle Dermatologie (LIED)*, Lübeck, Germany) for critical discussion and assistance in preparation of the manuscript.

REFERENCES

- Hammers CM, Stanley JR. Mechanisms of disease: pemphigus and bullous pemphigoid. *Annu Rev Pathol* (2016) 11:175–97. doi:10.1146/annurev-pathol-012615-044313
- Malik M, Ahmed AR. Concurrence of systemic lupus erythematosus and pemphigus: coincidence or correlation? *Dermatol Basel Switz* (2007) 214:231–9. doi:10.1159/000099588
- Calebotta A, Cirocco A, Giansante E, Reyes O. Systemic lupus erythematosus and pemphigus vulgaris: association or coincidence. *Lupus* (2004) 13:951–3. doi:10.1191/0961203304lu1073cr
- Rahman A, Isenberg DA. Systemic lupus erythematosus. *N Engl J Med* (2008) 358:929–39. doi:10.1056/NEJMra071297
- Prüßmann J, Prüßmann W, Recke A, Rentzsch K, Juhl D, Henschler R, et al. Co-occurrence of autoantibodies in healthy blood donors. *Exp Dermatol* (2014) 23:519–21. doi:10.1111/exd.12445
- Holtman IR, Raj DD, Miller JA, Schaafsma W, Yin Z, Brouwer N, et al. Induction of a common microglia gene expression signature by aging and neurodegenerative conditions: a co-expression meta-analysis. *Acta Neuropathol Commun* (2015) 3:31. doi:10.1186/s40478-015-0203-5
- Granlund Av, Flatberg A, Østvik AE, Drozdov I, Gustafsson BI, Kidd M, et al. Whole genome gene expression meta-analysis of inflammatory bowel disease colon mucosa demonstrates lack of major differences between Crohn's disease and ulcerative colitis. *PLoS One* (2013) 8(2):e56818. doi:10.1371/journal.pone.0056818
- Cárdenas-Roldán J, Rojas-Villarraga A, Anaya J-M. How do autoimmune diseases cluster in families? A systematic review and meta-analysis. *BMC Med* (2013) 11:73. doi:10.1186/1741-7015-11-73
- Troy NM, Hollams EM, Holt PG, Bosco A. Differential gene network analysis for the identification of asthma-associated therapeutic targets in allergen-specific T-helper memory responses. *BMC Med Genomics* (2016) 9:9. doi:10.1186/s12920-016-0171-z
- Zhao H, Cai W, Su S, Zhi D, Lu J, Liu S. Screening genes crucial for pediatric pilocytic astrocytoma using weighted gene coexpression network analysis combined with methylation data analysis. *Cancer Gene Ther* (2014) 21:448–55. doi:10.1038/cgt.2014.49
- Ring KL, An MC, Zhang N, O'Brien RN, Ramos EM, Gao F, et al. Genomic analysis reveals disruption of striatal neuronal development and therapeutic targets in human Huntington's disease neural stem cells. *Stem Cell Reports* (2015) 5:1023–38. doi:10.1016/j.stemcr.2015.11.005
- Nishifuji K, Amagai M, Kuwana M, Iwasaki T, Nishikawa T. Detection of antigen-specific B cells in patients with pemphigus vulgaris by enzyme-linked immunospot assay: requirement of T cell collaboration for autoantibody production. *J Invest Dermatol* (2000) 114:88–94. doi:10.1046/j.1523-1747.2000.00840.x
- Takahashi H, Kouno M, Nagao K, Wada N, Hata T, Nishimoto S, et al. Desmoglein 3-specific CD4+ T cells induce pemphigus vulgaris and interface dermatitis in mice. *J Clin Invest* (2011) 121:3677–88. doi:10.1172/JCI57379
- Mak A, Kow NY. The pathology of T cells in systemic lupus erythematosus. *J Immunol Res* (2014) 2014:e419029. doi:10.1155/2014/419029

FUNDING

This study was supported by the *Deutsche Forschungsgemeinschaft* through the training programs “Modulation of Autoimmunity” (grant number GRK 1727/1) and “Genes, Environment and Inflammation” (grant number GRK 1743/1).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at <http://www.frontiersin.org/articles/10.3389/fimmu.2017.01992/full#supplementary-material>.

- Konya C, Paz Z, Tsokos GC. The role of T cells in systemic lupus erythematosus: an update. *Curr Opin Rheumatol* (2014) 26:493–501. doi:10.1097/BOR.0000000000000082
- Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res* (2013) 41:D991–5. doi:10.1093/nar/gks1193
- Jeffries MA, Dozmorov M, Tang Y, Merrill JT, Wren JD, Sawalha AH. Genome-wide DNA methylation patterns in CD4+ T cells from patients with systemic lupus erythematosus. *Epigenetics* (2011) 6:593–601. doi:10.4161/epi.6.5.15374
- Malheiros D, Panepucci RA, Roselino AM, Araújo AG, Zago MA, Petzl-Erler ML. Genome-wide gene expression profiling reveals unsuspected molecular alterations in pemphigus foliaceus. *Immunology* (2014) 143:381–95. doi:10.1111/imm.12315
- Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* (2012) 28:882–3. doi:10.1093/bioinformatics/bts034
- Diez D, Alvarez R, Dopazo A. Codelink: an R package for analysis of GE healthcare gene expression bioarrays. *Bioinforma Oxf Engl* (2007) 23:1168–9. doi:10.1093/bioinformatics/btm072
- Gautier L, Cope L, Bolstad BM, Irizarry RA. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinforma Oxf Engl* (2004) 20:307–15. doi:10.1093/bioinformatics/btg405
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res* (2015) 43(7):e47. doi:10.1093/nar/gkv007
- Liu BH, Yu H, Tu K, Li C, Li YX, Li YY. DCGL: an R package for identifying differentially coexpressed genes and links from gene expression microarray data. *Bioinformatics* (2010) 26:2637–8. doi:10.1093/bioinformatics/btq471
- Miller JA, Cai C, Langfelder P, Geschwind DH, Kurian SM, Salomon DR, et al. Strategies for aggregating gene expression data: the collapseRows R function. *BMC Bioinformatics* (2011) 12:322. doi:10.1186/1471-2105-12-322
- Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostat Oxf Engl* (2007) 8:118–27. doi:10.1093/biostatistics/kxj037
- Dixon P. VEGAN, a package of R functions for community ecology. *J Veg Sci* (2003) 14:927–30. doi:10.1111/j.1654-1103.2003.tb02228.x
- Langfelder P, Horvath S. Eigengene networks for studying the relationships between co-expression modules. *BMC Syst Biol* (2007) 1:54. doi:10.1186/1752-0509-1-54
- Altay G, Emmert-Streib F. Inferring the conservative causal core of gene regulatory networks. *BMC Syst Biol* (2010) 4:132. doi:10.1186/1752-0509-4-132
- Xia J, Fjell CD, Mayer ML, Pena OM, Wishart DS, Hancock REW. INMEX—a web-based tool for integrative meta-analysis of expression data. *Nucleic Acids Res* (2013) 41:W63–70. doi:10.1093/nar/gkt338
- Thornton-Wells TA, Johnson KB. Perl Programming for biologists. *J Am Med Inform Assoc* (2004) 11:173. doi:10.1197/jamia.M1457
- Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* (2003) 4:3. doi:10.1186/gb-2003-4-5-p3

32. Xia J, Benner MJ, Hancock REW. Network Analyst--integrative approaches for protein-protein interaction network analysis and visual exploration. *Nucleic Acids Res* (2014) 42:W167-74. doi:10.1093/nar/gku443
33. Sarig O, Bercovici S, Zoller L, Goldberg I, Indelman M, Nahum S, et al. Population-specific association between a polymorphic variant in ST18, encoding a pro-apoptotic molecule, and pemphigus vulgaris. *J Invest Dermatol* (2012) 132:1798-805. doi:10.1038/jid.2012.46
34. Vodo D, Sarig O, Geller S, Ben-Asher E, Olender T, Bochner R, et al. Identification of a functional risk variant for pemphigus vulgaris in the ST18 gene. *PLoS Genet* (2016) 12:e1006008. doi:10.1371/journal.pgen.1006008
35. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature* (2009) 461:747-53. doi:10.1038/nature08494
36. Cojocaru M, Cojocaru IM, Silosi I. Multiple autoimmune syndrome. *Maedica (Buchar)* (2010) 5:132-4.
37. Sawamura S, Kajihara I, Makino K, Makino T, Fukushima S, Jinnin M, et al. Systemic lupus erythematosus associated with myasthenia gravis, pemphigus foliaceus and chronic thyroiditis after thymectomy. *Australas J Dermatol* (2016) 58(3):e120-2. doi:10.1111/ajd.12510
38. Theofilopoulos AN, Baccala R, Beutler B, Kono DH. Type I interferons (alpha/beta) in immunity and autoimmunity. *Annu Rev Immunol* (2005) 23:307-36. doi:10.1146/annurev.immunol.23.021704.115843
39. Olson JK, Croxford JL, Miller SD. Virus-induced autoimmunity: potential role of viruses in initiation, perpetuation, and progression of T-cell-mediated autoimmune disease. *Viral Immunol* (2001) 14:227-50. doi:10.1089/088282401753266756
40. Phillips PE. The virus hypothesis in systemic lupus erythematosus. *Ann Intern Med* (1975) 83:709-15. doi:10.7326/0003-4819-83-5-709
41. Denman AM. Systemic lupus erythematosus--is a viral aetiology a credible hypothesis? *J Infect* (2000) 40:229-33. doi:10.1053/jinf.2000.0670
42. Ramos-Casals M. Viruses and lupus: the viral hypothesis. *Lupus* (2008) 17:163-5. doi:10.1177/0961203307086268
43. Krain LS. Pemphigus. Epidemiologic and survival characteristics of 59 patients, 1955-1973. *Arch Dermatol* (1974) 110:862-5. doi:10.1001/archderm.1974.01630120012002
44. Marzano AV, Turlaki A, Merlo V, Spinelli D, Venegoni L, Crosti C. Herpes simplex virus infection and pemphigus. *Int J Immunopathol Pharmacol* (2009) 22:781-6. doi:10.1177/039463200902200324
45. Senger P, Sinha AA. Exploring the link between herpes viruses and pemphigus vulgaris: literature review and commentary. *Eur J Dermatol* (2012) 22:728-35. doi:10.1684/ejd.2012.1836
46. Takahashi I, Kobayashi TK, Suzuki H, Nakamura S, Tezuka F. Coexistence of pemphigus vulgaris and herpes simplex virus infection in oral mucosa diagnosed by cytology, immunohistochemistry, and polymerase chain reaction. *Diagn Cytopathol* (1998) 19:446-50. doi:10.1002/(SICI)1097-0339(199812)19:6<446::AID-DC8>3.0.CO;2-2
47. Kurata M, Mizukawa Y, Aoyama Y, Shiohara T. Herpes simplex virus reactivation as a trigger of mucous lesions in pemphigus vulgaris. *Br J Dermatol* (2014) 171:554-60. doi:10.1111/bjd.12961
48. Seo J-Y, Yaneva R, Cresswell P. Viperin: a multifunctional, interferon-inducible protein that regulates virus replication. *Cell Host Microbe* (2011) 10:534-9. doi:10.1016/j.chom.2011.11.004
49. Becker AM, Dao KH, Han BK, Kornu R, Lakhanpal S, Mobley AB, et al. SLE peripheral blood B cell, T cell and myeloid cell transcriptomes display unique profiles and each subset contributes to the interferon signature. *PLoS One* (2013) 8:e67003. doi:10.1371/journal.pone.0067003
50. Gan L, O'Hanlon TP, Lai Z, Fannin R, Weller ML, Rider LG, et al. Gene expression profiles from disease discordant twins suggest shared antiviral pathways and viral exposures among multiple systemic autoimmune diseases. *PLoS One* (2015) 10(11):e0142486. doi:10.1371/journal.pone.0142486
51. Tischner D, Woess C, Ottina E, Villunger A. Bcl-2-regulated cell death signalling in the prevention of autoimmunity. *Cell Death Dis* (2010) 1:e48. doi:10.1038/cddis.2010.27
52. Choi UY, Kang J-S, Hwang YS, Kim Y-J. Oligoadenylate synthase-like (OASL) proteins: dual functions and associations with diseases. *Exp Mol Med* (2015) 47:e144. doi:10.1038/emmm.2014.110
53. Ferreira RC, Guo H, Coulson RM, Smyth DJ, Pekalski ML, Burren OS, et al. A type I interferon transcriptional signature precedes autoimmunity in children genetically at risk for type 1 diabetes. *Diabetes* (2014) 63:2538-50. doi:10.2337/db13-1777
54. Wang D, Qin H, Du W, Shen YW, Lee WH, Riggs AD, et al. Inhibition of S-phase kinase-associated protein 2 (Skp2) reprograms and converts diabetogenic T cells to Foxp3+ regulatory T cells. *Proc Natl Acad Sci U S A* (2012) 109:9493-8. doi:10.1073/pnas.1207293109
55. Hunziker T, Nydegger UE, Lerch PG, Vassalli JD. Platelet-derived factors enhance pemphigus acantholysis in skin organ cultures. *Clin Exp Immunol* (1986) 64:442-9.
56. Mizutani H, Ohyanagi S, Nouchi N, Inachi S, Shimizu M. Tissue factor and thrombomodulin expression on keratinocytes as coagulation/anti-coagulation cofactor and differentiation marker. *Australas J Dermatol* (1996) 37(Suppl):1. doi:10.1111/j.1440-0960.1996.tb01085.x
57. Chrabot BS, Kariuki SN, Zervou MI, Feng X, Arrington J, Jolly M, et al. Genetic variation near IRF8 is associated with serologic and cytokine profiles in systemic lupus erythematosus and multiple sclerosis. *Genes Immun* (2013) 14:471-8. doi:10.1038/gene.2013.42
58. Choi J-M, Bothwell ALM. The nuclear receptor PPARs as important regulators of T-cell functions and autoimmune diseases. *Mol Cells* (2012) 33:217-22. doi:10.1007/s10059-012-2297-y
59. Zhu J, Jankovic D, Grinberg A, Guo L, Paul WE. Gfi-1 plays an important role in IL-2-mediated Th2 cell expansion. *Proc Natl Acad Sci U S A* (2006) 103:18214-9. doi:10.1073/pnas.0608981103
60. McCarthy FP, Delany AC, Kenny LC, Walsh SK. PPAR- γ --a possible drug target for complicated pregnancies. *Br J Pharmacol* (2013) 168:1074-85. doi:10.1111/bph.12069

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Sezin, Vorobyev, Sadik, Zillikens, Gupta and Ludwig. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.