Cross-Comparison of Inflammatory Skin Disease Transcriptomics Identifies PTEN as a Pathogenic Disease Classifier in Cutaneous Lupus Erythematosus

JID Open

Brian D. Aevermann[1,2,17], Jeremy Di Domizio[3,17], Peter Olah[4], Fanny Saidoune[3], John M. Armstrong[5], Hervé Bachelez[6], Jonathan Barker[7], Muzlifah Haniffa[8], Valerie Julia[5], Kasper Juul[9], Jayendra Kumar Krishnaswamy[5], Thomas Litman[9], Ian Parsons[10], Kavita Y. Sarin[11], Matthias Schmuth[12], Michael Sierra[9], Michael Simpson[13], Bernhard Homey[4], Christopher E.M. Griffiths[14], Richard H. Scheuermann[1,15,16] and Michel Gilliet[3]

Tissue transcriptomics is used to uncover molecular dysregulations underlying diseases. However, the majority of transcriptomics studies focus on single diseases with limited relevance for understanding the molecular relationship between diseases or for identifying disease-specific markers. In this study, we used a normalization approach to compare gene expression across nine inflammatory skin diseases. The normalized datasets were found to retain differential expression signals that allowed unsupervised disease clustering and identification of disease-specific gene signatures. Using the NS-Forest algorithm, we identified a minimal set of biomarkers and validated their use as diagnostic disease classifier. Among them, PTEN was identified as being a specific marker for cutaneous lupus erythematosus and found to be strongly expressed by lesional keratinocytes in association with pathogenic type I IFNs. In fact, PTEN facilitated the expression of IFN-β and IFN-κ in keratinocytes by promoting activation and nuclear translocation of IRF3. Thus, cross-comparison of tissue transcriptomics is a valid strategy to establish a molecular disease classification and to identify pathogenic disease biomarkers.

Journal of Investigative Dermatology (2024) **144,** 252–262; doi:10.1016/j.jid.2023.06.211

[1]*J. Craig Venter Institute, La Jolla, California, USA;* [2]*Chan Zuckerberg Initiative, Redwood City, California, USA;* [3]*Department of Dermatology, Lausanne University Hospital (CHUV), University of Lausanne, Lausanne, Switzerland;* [4]*Department of Dermatology, Medical Faculty, University Hospital Düsseldorf, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany;* [5]*Galderma, Lausanne, Switzerland;* [6]*Department of Dermatology, Paris Diderot University, Paris, France;* [7]*St John's Institute of Dermatology, Faculty of Life Sciences & Medicine, Kings College London, London, United Kingdom;* [8]*Department of Dermatology and NIHR Newcastle Biomedical Research Centre, Newcastle, United Kingdom;* [9]*LEO Pharma A/S, Ballerup, Denmark;* [10]*Celgene International Sarl, Boudry, Switzerland;* [11]*Department of Dermatology, Stanford University Medical Center, Palo Alto, California, USA;* [12]*Department of Dermatology, Venereology and Allergy, Medical University Innsbruck, Innsbruck, Austria;* [13]*Department of Genomic Medicine, King's College London, London, United Kingdom;* [14]*Section of Dermatology, Department of Medicine, University of Manchester, Manchester, United Kingdom;* [15]*Department of Pathology, University of California San Diego School of Medicine, La Jolla, California, USA; and* [16]*La Jolla Institute for Immunology, La Jolla, California, USA*

[17]*These authors contributed equally to this work.*

Correspondence: Michel Gilliet, Department of Dermatology, Lausanne University Hospital (CHUV), University of Lausanne, Avenue de Beaumont 29, Lausanne 1011, Switzerland. E-mail: michel.gilliet@chuv.ch

Abbreviations: AA, alopecia areata; AD, atopic dermatitis; Akt, protein kinase B; CLE, cutaneous lupus erythematosus; DE, differential expression; HS, hidradenitis suppurativa; KC, keratinocyte; LE, lupus erythematosus; LP, lichen planus; pDC, plasmacytoid dendritic cell; PI3K, phosphatidylinositol 3 kinase; PIP3, phosphatidylinositol-3, 4, 5-triphosphate; RMA, Robust Multi-chip Average; siRNA, small interfering RNA

Received 16 August 2022; revised 5 June 2023; accepted 9 June 2023; accepted manuscript published online 19 August 2023; corrected proof published online 30 September 2023

INTRODUCTION

Genomics is revolutionizing the medical sciences, and skin science has been at the forefront of applied molecular-based treatments. Such therapies have been successfully applied to major inflammatory skin diseases, such as psoriasis (Conrad and Gilliet, 2018; Griffiths et al., 2021), atopic dermatitis (AD) (Bieber, 2022; Puar et al., 2021), and lupus erythematosus (LE) (Chasset and Francès, 2019; Wenzel, 2019), and are currently moving to address other skin diseases, including hidradenitis suppurativa (HS) (Goldburg et al., 2020), alopecia areata (AA) (Wang et al., 2018), and vitiligo (Frisoli et al., 2020). Despite recent successes, the underlying molecular mechanisms of many skin diseases are still unknown, and new, more precise molecular targets are needed for both diagnostic and therapeutic applications. Tissue transcriptomics has been the leading methodology for investigating the molecular basis of diseases (Banchereau et al., 2017), including those of the skin. Investigators in skin science may have an edge over researchers of other organs because the skin is easily accessible and routinely biopsied. As a result, many skin transcriptomics datasets are available in public repositories for analysis, including the Gene Expression Omnibus by the National Center for Biotechnology Infomation and ArrayExpress by the European Bioinformatics Institute.

Although the rate of data generation and investigation of skin diseases has led to great advances in the field, the

majority of studies have focused on single diseases and their controls. Only a few studies have attempted a systematic molecular comparison between multiple diseases to identify gene signatures or molecular biomarkers that are disease specific (Inkeles et al., 2015; Mo et al., 2018; Wang et al., 2016). We hypothesized that the identification of such disease-specific molecular markers through a comparative approach of tissue transcriptomic involving multiple diseases would provide the basis for molecular classification of skin diseases as a diagnostic tool and would unravel disease-specific mechanisms relevant for disease pathogenesis, which can be further developed as therapeutic targets.

We compared the transcriptional profiles of nine inflammatory skin diseases, including psoriasis, AD, LE, rosacea, vitiligo, AA, acne, HS, and lichen planus (LP), obtained from independent studies to identify disease-specific gene expression biomarkers. The obtained biomarkers were further validated to produce a molecular classification of these diseases using independent datasets. Among the markers, we identified PTEN as being a cutaneous lupus erythematosus (CLE) classifier and further investigated its functional implications in disease pathogenesis. PTEN was strongly expressed and activated in lesional keratinocytes (KCs) of CLE and was found to promote pathogenic type I IFN expression by promoting nuclear translocation of phosphorylated IRF3. Thus, comparison of transcriptional profiles of multiple diseases from independent studies is a powerful tool to generate a molecular disease cartography and unravel pathomechanism through the identification of disease-specific markers.

RESULTS

Normalized transcriptomics data retain differential gene expression signals and cluster by disease

To compare gene expression across multiple inflammatory skin diseases and determine their gene expression differences, we mined public repositories and selected nine microarray datasets on the basis of the sample size and the availability of healthy skin control samples: psoriasis (Fyhrquist et al., 2019), AD (Fyhrquist et al., 2019), LP (Shao et al., 2019), CLE (Liu et al., 2017), rosacea (Buhl et al., 2015), vitiligo (Regazzetti et al., 2015), AA (Jabbari et al., 2016), HS (Blok et al., 2016), and acne vulgaris (acne) (Kelhälä et al., 2014). When comparing expression patterns for individual genes between these experiments, both the disease samples and the control samples showed different expression ranges between experiments (Figure 1a). For example, the $\log_2$ mean expression of *JAK2* in the control samples in the LP experiment was ~0, whereas in the CLE experiment, the mean expression in the controls was ~3. Consequently, we performed cross-experiment normalization in a manner similar in concept to batch correction, whereby each experiment is transformed into a comparable data range, assuming that gene expression in healthy control samples is similar across experiments.

There are many methods available for microarray experiment normalization, the most common of which is quantile normalization using Robust Multichip Average (RMA) (e.g., in the Limma package). However, these normalization strategies are designed for within-experiment normalization in which technical noise is estimated across all samples and removed (Irizarry et al., 2003). In all cases, the publicly available data had already been processed by one of these within-experiment normalization methods. To achieve cross-experiment normalization, we applied RatioA normalization to each experiment, subtracting the mean gene expression level for each gene in all healthy control samples in a given experiment and applying a $\log_2$ transformation, thereby centering the overall expression distribution in healthy samples around zero (e.g., for both the LP and CLE experiments) (Figure 1b). These gene-specific normalization factors are then applied to both healthy and disease samples to generate the postnormalized dataset. Importantly, whereas the gene expression distributions in the control samples were similar after normalization, the strong *JAK2* upregulation in LP samples and the modest upregulation in CLE samples observed in the original prenormalized dataset were preserved in the postnormalized dataset (Figure 1a).

If done appropriately, cross-experiment normalization should not affect the original biological signal. To evaluate the impact of normalization on the gene expression signal, we performed differential expression (DE) analysis between samples from individuals with the disease and those from healthy control on both prenormalization and postnormalization datasets. DE sets prior to normalization were compared with the previously published results, and no major discrepancies were found. After filtering the DE results to select genes with an adjusted $P < 0.05$, the intersection of the DE results was computed as a measure of signal retention (Figure 1c). Overall, there was high preservation of differentially expressed signal, with a median Jaccard index of 0.84. The lowest overlap was found with the AA study (GSE68801), in which the Jaccard index was 0.54. Comparisons were not performed for vitiligo owing to the low number of differentially expressed genes in the original experiment, which may reflect a subtle biological signal for vitiligo from biopsies in nonactive disease.

The data were then fully integrated by merging all eight normalized gene expression datasets on the basis of the genes shared between datasets. The final integrated normalized gene expression dataset contained 478 disease and 365 control samples and 14,673 common genes.

Unsupervised clustering was performed using the integrated dataset and a combination of k means and hierarchical clustering as implemented in the SC3 package, with optimal k determined using silhouette scoring (see Materials and Methods). The disease phenotypes for each sample, obtained from the sample metadata provided with each experiment, were enriched in the distinct k-means clusters (Figure 2a). Cluster #1 contained all the CLE samples, cluster #2 contained rosacea samples, cluster #3 contained AA samples, cluster #4 contained LP samples, cluster #5 contained vitiligo samples, cluster #6 contained psoriasis samples, cluster #7 contained AD samples, cluster #8 contained both acne and HS samples, and cluster #9 contained the vast majority of control samples. The clustering of control samples from all experiments into one major cluster, as opposed to the disease samples from the same experiment, supports the conclusion that the cross-experiment normalization was a success and that the difference detected between diseases should not be due to batch effects. Of the nine sample clusters generated,

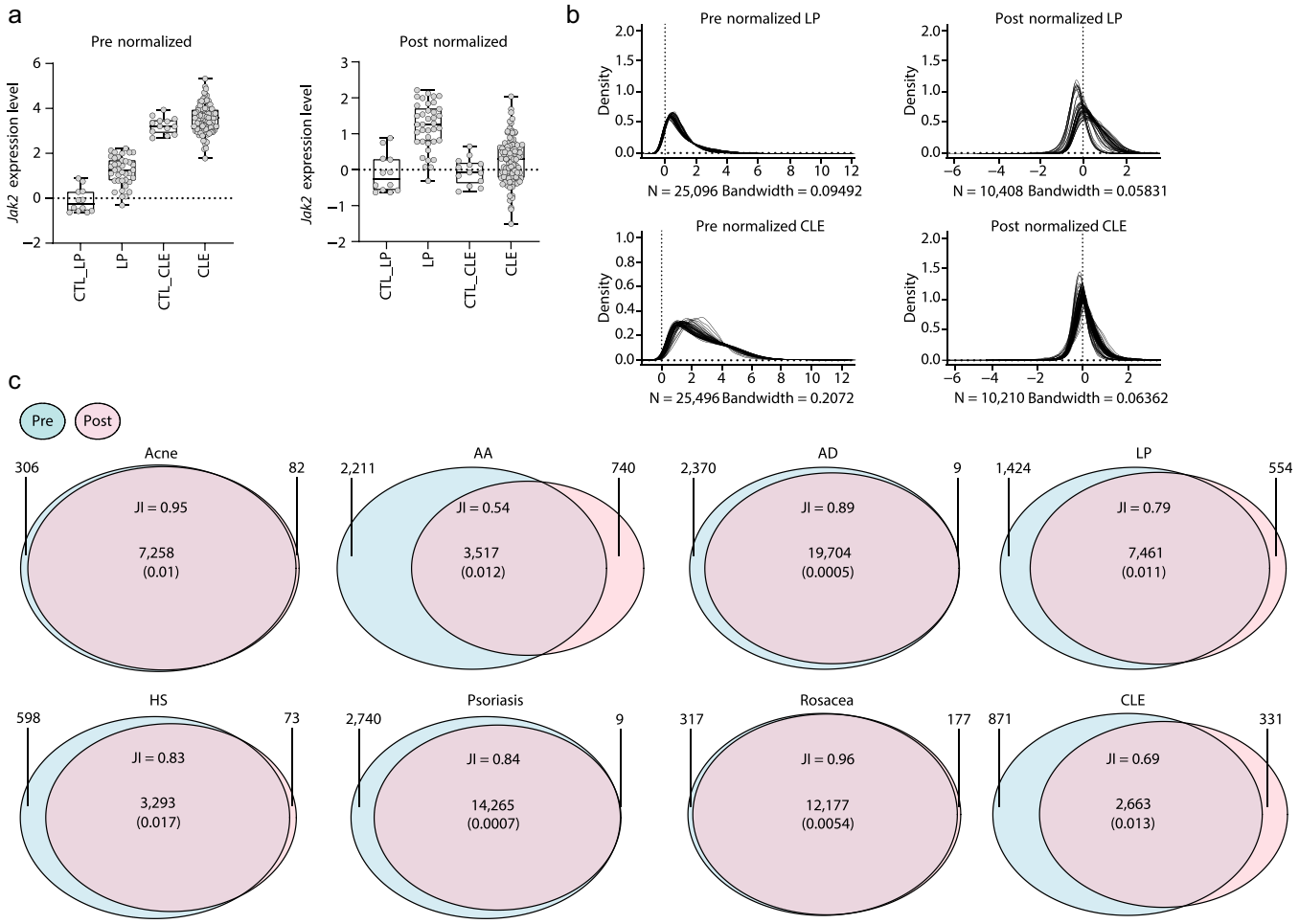


Figure 1. Normalization approach for integrative meta-analysis across microarray disease experiments. (**a**) As an example, prenormalized expression values for *JAK2* are distributed across different scales in different experiments, reflected by the different CTL sample distributions (left). In contrast, postnormalized *JAK2* values show similar distributions in CTL samples in the two experiments while retaining the disease-specific differential expression observed before normalization (right). Expression levels of *JAK2* in healthy (CTL) and lesional skin of patients with LP and CLE were used for illustration. (**b**) Distribution of expression data across all genes of the LP (top) and CLE (bottom) samples before (left) and after (right) normalization. (**c**) Differential expression comparison—the genes that were found to be differentially expressed by Limma analysis both before or after normalization were compared in the different skin disease groups, and a Venn diagram showing their overlap is presented. The number of genes before and after normalization; the JI, calculated as a measure of their similarity; and the *P*-value of overlap, indicated in parentheses, calculated using a Wilcoxon rank sum statistical test are shown. AD, atopic dermatitis; CLE, cutaneous lupus erythematosus; CTL, control; JI, Jaccard Index; LP, lichen planus.

only two clusters had a mixture of sample phenotypes: cluster #3 contains samples only from the AA experiment but a number of control samples mixed with the disease samples (23% of the cluster consists of control samples), potentially related to a unique transcriptional scalp microenvironment, and cluster #8 contains the disease samples from both acne and HS, potentially reflecting the pathogenic similarities with dominant neutrophilic infiltration.

The hierarchical clustering results are consistent with the k-means clustering and provide additional insights regarding the relationships between these inflammatory skin diseases at the transcriptional level. For example, LP, LE, AA, and vitiligo share a common cluster branch, indicating that they are more similar to each other than to other skin diseases. This finding is in line with the common cytotoxic type I and II IFN–dominated transcriptional profiles with a lichenoid reaction pattern described for these diseases (Eyerich and Eyerich, 2018). Psoriasis and AD also share a common branch, reflecting the high degree of genetic overlap of these diseases, despite distinct clinical phenotypes and opposing immune mechanisms (Baurecht et al., 2015). Surprisingly, rosacea appears more transcriptionally related to type I and II IFN–dominated diseases LE and LP than to acne, a finding potentially related to the recent finding of a role of type I IFNs in rosacea (Mylonas et al., 2023).

Identification and validation of disease classifiers

To identify disease-specific molecular markers, we used the NS-Forest algorithm. As the first approach, we obtained an extended list of markers that showed a binary expression pattern (in which a gene is expressed in the targeted disease while being absent in the other diseases) (Figure 2b and Supplementary Table S1). For each of the diseases in the molecular taxonomy, several genes with highly binary expression patterns were found, many of which have been previously reported, such as *NOS2* (Quaranta et al., 2014),

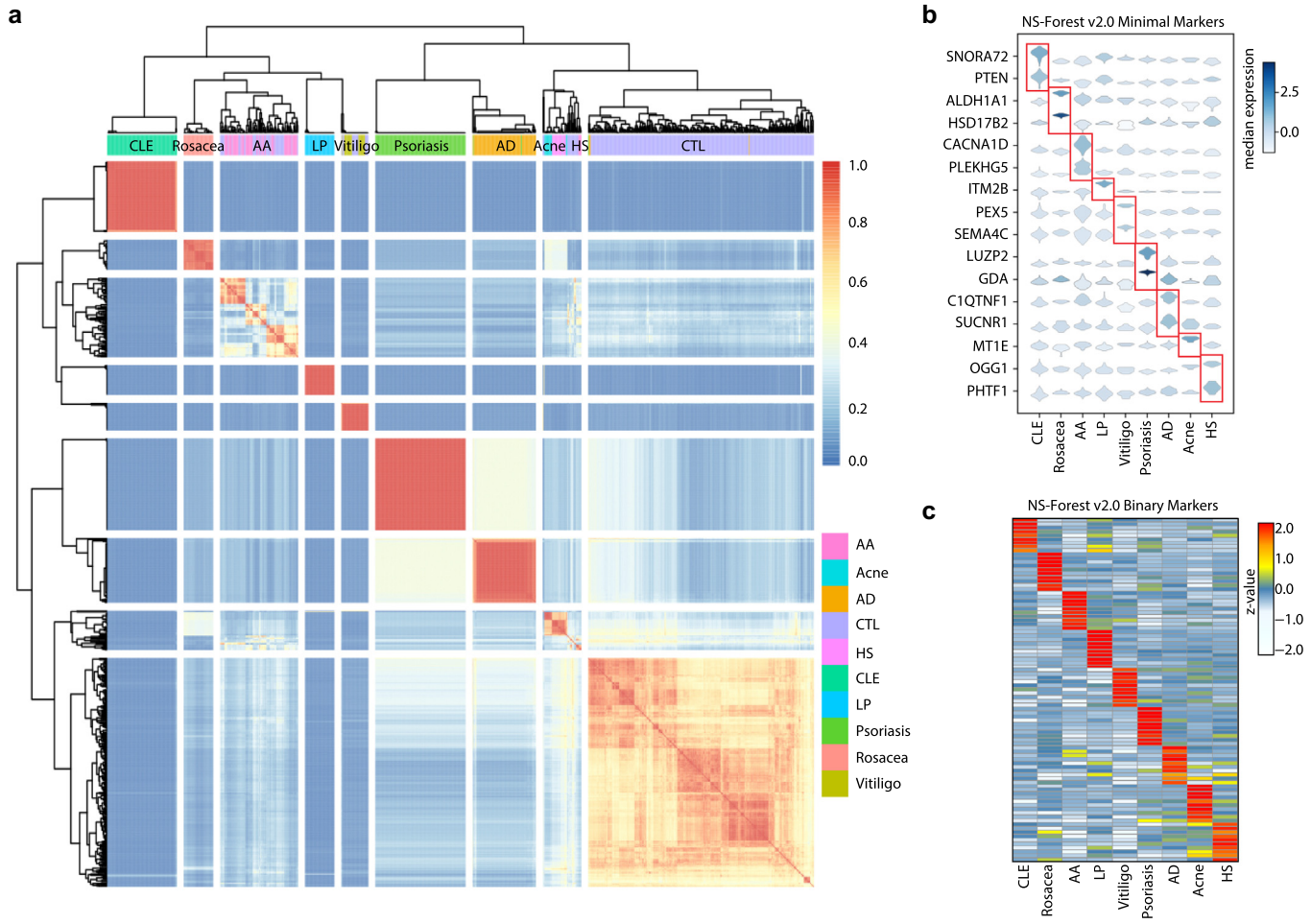


Figure 2. Unsupervised hierarchical clustering of skin disease transcriptional profiles. (**a**) CTL RatioA-normalized gene expression data of the different skin disease groups, including healthy skin samples (CTL, n = 329) and lesional skin samples of plaque-type psoriasis vulgaris (psoriasis, n = 120), AD (n = 83), LP (n = 38), CLE (n = 90), HS (n = 17), acne vulgaris (acne, n = 12), rosacea (n = 38), AA (n = 60), and vitiligo (n = 20) were merged on the intersection of shared genes and then clustered using the SC3 package. Heatmap is colored by the frequency of a sample being within a cluster estimated by the aggregate of ensembled cluster solutions for that k. (**b**) Violin expression plots of NS-Forest, version 2.0, minimal markers for each disease group colored by median expression in lesional samples. (**c**) Heatmap of extended binary markers for each disease group from NS-Forest colored by median expression in lesional samples. AA, alopecia areata; AD, atopic dermatitis; CLE, cutaneous lupus erythematosus; CTL, control; HS, hidradenitis suppurativa; LP, lichen planus.

PLA2G4D (Shao et al., 2021), and *VNN3* (Jansen et al., 2009) in psoriasis. To validate our approach, we took the top five psoriasis and AD markers from the extended NS-Forest marker list and tested their ability to classify independent transcriptomic datasets, generated with a different technology (RNA sequencing). F-beta scores were very high (0.76−1.00) in these independent datasets (Supplementary Figure S1), indicating that the markers determined from the microarray-based datasets have high precision and recall in independent experiments, even using a distinct technological platform.

We then used the NS-Forest algorithm to find minimal disease classifiers. The median number of markers per disease was two (Figure 2c) and included *CACNA1D* and *PLEKHG5* for AA, *MT1E* for acne, *C1QTNF1* and *SUCNR1* for AD, *OGG1* and *PHTF1* for HS, *PTEN* and *SNORA72* for CLE, *ITM2B* for LP, *GDA* and *LUZP2* for psoriasis, *ALDH1A1* and *HSD17B2* for rosacea, and *PEX5* and *SEMA4C* for vitiligo. The use of these minimal markers for classification showed very high precision with a median F-beta score of 0.96 (range = 0.87−1.00) (Table 1). Thus, these markers are truly disease specific and are reproducibly diagnostic. Interestingly, the majority of these markers have not been previously reported as differentially expressed genes in single skin disease studies.

PTEN is a CLE biomarker that is overexpressed by KCs

Having identified minimal markers that reproducibly classify inflammatory skin diseases, we next sought to determine whether they have functional relevance and are part of pathogenic mechanisms in the specific diseases. We selected *PTEN* (phosphatase and TENsin homolog), a tumor suppressor gene and negative regulator of phosphatidylinositol 3 kinase (PI3K)/PIP3 signaling pathway, for an in-depth evaluation. *PTEN* was identified as a classifier for CLE (Figure 2b), and its mRNA expression was significantly higher in CLE lesions than in all other inflammatory skin diseases or in healthy skin (ANOVA $F = 78$, $P < 0.0001$) (Figure 3a). Immunohistochemistry revealed that PTEN protein expression was particularly strong in the epidermis of CLE compared with that in the epidermis of

Table 1. NS-Forest Markers

Disease	f-Beta	Precision	Recall	True Negative	False Positive	False Negative	True Positive	Marker Count	Marker Gene1	Marker Gene2
AA	0.89	1.00	0.62	418	0	23	37	2	*CACNA1D* ≥ 0.52	*PLEKHG5* ≥ 0.21
Acne	0.96	1.00	0.83	466	0	2	10	1	*MT1E* ≥ 1.24	
AD	0.86	0.98	0.58	394	1	35	48	2	*C1QTNF1* ≥ 0.76	*SUCNR1* ≥ 0.57
HS	0.88	1.00	0.59	461	0	7	10	2	*OGG1* ≥ 0.39	*PHTF1* ≥ 0.91
CLE	0.86	0.96	0.61	386	2	35	55	2	*PTEN* ≥ 0.61	*SNORA72* ≥ 0.79
LP	1.00	1.00	1.00	440	0	0	38	1	*ITM2B* ≥ 0.73	
PSO	1.00	1.00	0.98	358	0	2	118	2	*GDA* ≥ 3.16	*LUZP2* ≥ 0.69
Rosacea	0.98	1.00	0.89	440	0	4	34	2	*ALDH1A1* ≥ 1.19	*HSD17B2* ≥ 2.96
Vitiligo	0.97	1.00	0.85	458	0	3	17	2	*PEX5* ≥ 0.52	*SEMA4C* ≥ 0.44

Abbreviations: AA, alopecia areata; AD, atopic dermatitis; CLE, cutaneous lupus erythematosus; HS, hidradenitis suppurativa; LP, lichen planus; PSO, psoriasis.

The minimum set of marker genes identified by the NS-Forest algorithm for each of the nine inflammatory skin diseases is listed. The index column includes the disease abbreviation, marker gene(s), and the expression threshold(s) determined by NS-Forest and used for classification. The F-beta score, true negative, false positive, false negative, and true positive values across all 478 lesional samples using the combined gene sets are listed. The sum of true positive and false negative values gives the number of samples evaluated for each skin disease.

other inflammatory skin diseases or healthy skin (ANOVA F = 2.3, $P < 0.05$) (Figure 3b).

We then sought to determine whether the increased PTEN expression in CLE has functional activity by negatively regulating PIP3-dependent processes (De Marco et al., 2017). Indeed, expression of a subset of PIP3 downstream genes—*ASNS*, *PSATI*, *SDSL*, and *ASS1*—was significantly decreased in CLE lesions compared with the expression in all other inflammatory skin diseases or healthy skin, as shown in a composite score of PIP3-dependent genes (ANOVA F = 41, $P < 0.0001$) (Figure 3c). Together, these data indicate that the CLE classifier PTEN is strongly overexpressed in CLE epidermis and exerts functional activities.

PTEN promotes nuclear IRF3 translocation in KCs with induction of pathogenic type I IFNs in CLE

LE, including CLE, is a type I IFN–driven disease (Banchereau et al., 2016; Blanco et al., 2001; Garcia-Romo et al., 2011; Lande et al., 2011; Rönnblom et al., 2003; Vallin et al., 1999), as recently evidenced by the therapeutic

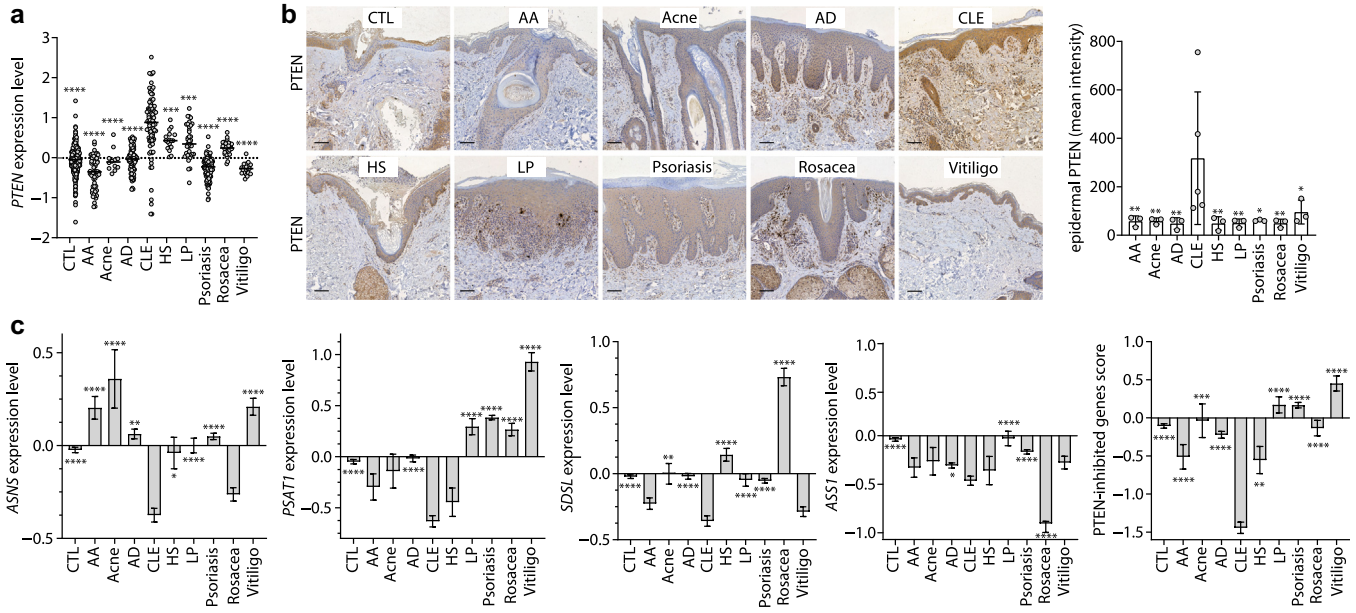


Figure 3. Increased expression and activation of PTEN in keratinocytes of CLE. (**a**) CTL RatioA expression level of *PTEN* in the different skin disease groups using the SSF Bioinformatics Hub. (**b**) Immunohistochemical staining of PTEN in skin sections from the different patients (left) and quantification of its intensity in the epidermis (right). Bars = 100 mm. (**c**) CTL RatioA expression levels of the PTEN-inhibited target genes *ASNS*, *PSAT1*, *SDSL*, and *ASS1* as well as a cumulative score in healthy (CTL) and lesional skin from the different patients. (**a**, **c**) Data were statistically analyzed using Brown–Forsythe and Welch ANOVA followed by Dunnett's T3 multiple comparisons tests. *$P < 0.05$, **$P < 0.005$, ***$P < 0.0005$, and ****$P < 0.0001$. CTL, n = 329; AA, n = 60; acne, n = 12; AD, n = 83; CLE, n = 90; HS, n = 17; LP, n = 38; PSO, n = 120; rosacea, n = 38; and vitiligo, n = 20. (**b**) Data were statistically analyzed using one-way ANOVA followed by uncorrected Fisher's LSD comparisons test. *$P < 0.05$ and **$P < 0.005$. Error bars are SEM. CLE, n = 4; other diseases, n = 3. AA, alopecia areata; AD, atopic dermatitis; CLE, cutaneous lupus erythematosus; CTL, control; HS, hidradenitis suppurativa; LP, lichen planus; LSD, least significant difference; PSO, psoriasis; SSF, Skin Science Foundation.

efficacy of anti-IFNAR antibodies (Morand et al., 2020). Accordingly, we observed a strong type I IFN signature with expression of several IFN-stimulated genes in CLE, whereas other skin diseases showed either weak expression (psoriasis and LP) (Nestle et al., 2005) or no expression at all (Figure 4a). We then plotted the type I IFN signature in CLE against the expression of *PTEN* mRNA and found a highly significant correlation (Pearson r = 0.56, $P < 0.0001$) (Figure 4b), suggesting that the PTEN overexpression in CLE may be linked to the induction of pathogenic type I IFNs.

Although plasmacytoid dendritic cells (pDCs) represent the principal IFN-α producers in CLE (Farkas et al., 2001), recent studies point to KCs as another major source of type I IFNs in CLE through the production of IFN-β and the KC-specific IFN-κ (Sarkar et al., 2018; Scholtissek et al., 2017). In fact, immunofluorescence staining revealed a strong expression of both IFN-β and IFN-κ in CLE skin lesions with a predominant expression by KCs (unpaired *t*-test, $h^2 = 0.55$, $P < 0.0001$) (Figure 4c and Supplementary Figures S2 and S3). The IFN-β and IFN-κ expression by KCs directly correlated with epidermal PTEN protein expression (Pearson r = 0.65 and 0.73 respectively, $P < 0.0001$) (Figure 4d), supporting the hypothesis that PTEN overexpression may drive type I IFN production in CLE KCs.

During viral infections, PTEN has been linked to antiviral type I IFN responses in monocytes by promoting the

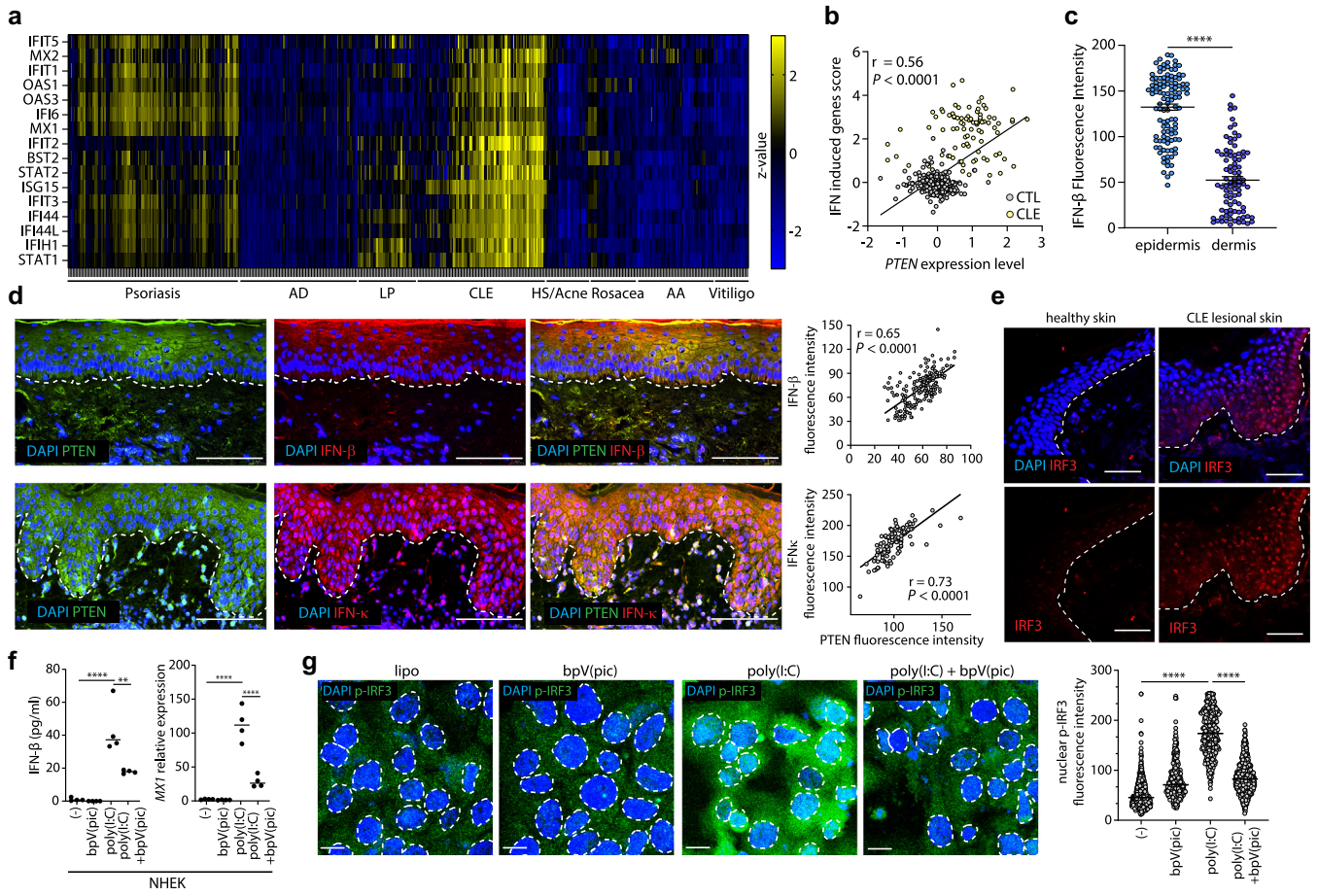


Figure 4. PTEN drives pathogenic type I IFN in CLE. (**a**) Heatmap showing control RatioA-normalized expression of selected IFN-response genes in the lesional skin of the different patients. Color intensity represents the level of normalized gene expression. (**b**) Correlation between the IFN-induced genes scores, calculated as the mean expression of the IFN-induced genes, and the expression level of PTEN in healthy and lesional skin of patients with CLE. Each dot represents a single patient. (**c**) Quantification of IFN-β fluorescence intensity at the single-cell level in epidermal and dermal cells of a lesional CLE skin specimen. (**d**) Confocal microscopy images of a representative lesional CLE skin (left) stained for PTEN (green), IFN-β (top) or IFN-k (bottom) (red), and DNA (blue). The dashed line delineates the dermo–epidermal junction. Bars = 100 mm. Correlation between PTEN and IFN-β or IFN-k fluorescence intensities at the single-cell level is shown (n = 177 and n = 100, respectively) (right). (**e**) Confocal microscopy images of representative healthy (left) and CLE lesional skin (right) stained for IRF3 (red) and DNA (blue). The dashed line delineates the dermo–epidermal junction. Bars = 50 mm. (**f**) Production of IFN-β and relative expression of MX1 in NHEKs stimulated with empty liposomes (−) or liposomes containing poly(I:C) in the presence or absence of the PTEN inhibitor bpV(pic). (**g**) Representative confocal microscopy images of NHEK (left) stimulated as in **d** and stained for p-IRF3 (green). Dashed lines delineate the nuclei. Bars = 20 mm. Quantification of nuclear p-IRF3 fluorescence intensities at the single-cell level is shown (n = 650) (right). (**b, d**) Pearson correlation coefficient and a two-tailed statistical significance are provided. (**c**) Bars represent the mean ± SEM. Data were statistically analyzed using unpaired *t*-test. ****$P < 0.0001$. (**e, f**) Representative data from one of four independent experiments are shown. Bars represent the means of four replicates. Data were statistically analyzed using one-way ANOVA followed by Tukey's multiple comparisons test. *$P < 0.05$, ***$P < 0.001$, and ****$P < 0.0001$. AA, alopecia areata; AD, atopic dermatitis; HS, hidradenitis suppurativa; CLE, cutaneous lupus erythematosus; NHEK, normal human epidermal keratinocyte; LP, lichen planus; p-IRF3, phosphorylated IRF3; PSO, psoriasis.

phosphorylation and nuclear translocation of IRF3 (Li et al., 2016). In CLE, epidermal type I IFN responses are also likely to involve IRF3 because it represents the master signaling molecule downstream of cytosolic nucleic acid sensors, such as RIG-I/MAVS and STING (Scholtissek et al., 2017). Immunofluorescence staining of CLE sections indeed revealed strong IRF3 staining with predominant nuclear localization in KCs, indicating that IRF3 was phosphorylated and translocated into the nucleus, leading to IFN responses in CLE KCs (Figure 4e and Supplementary Figure S4).

To investigate the involvement of PTEN, we stimulated normal human epidermal KCs with the poly(I:C), a RIG-I/MAVS agonist in KCs (Min et al., 2022) that triggers IRF3 activation and type I IFN responses with or without the addition of the PTEN inhibitor bpV(pic) (Schmid et al., 2004). RIG-I/MAVS activation of KCs induced the production of IFN-β and expression of the IFN-stimulated gene *MXA*, which were largely abrogated in the presence of the PTEN inhibitor (ANOVA $F = 27$ and 62, respectively, $P < 0.0001$) (Figure 4f). RIG-I/MAVS activation also induced cytosolic marked nuclear translocation of IRF3, a process that was strongly inhibited in the presence of the PTEN inhibitor (ANOVA $F = 975$, $P < 0.0001$) (Figure 4g). To confirm these data, we also performed small interfering RNA (siRNA)—mediated transient PTEN knockdown in KC and showed inhibition of IRF3 nuclear translocation in response to poly(I:C) stimulation (Supplementary Figure S5).

Taken together, these data indicate that the CLE classifier PTEN drives IRF3 signaling in CLE KCs, leading to pathogenic type I IFN responses by its ability to promote IRF3 phosphorylation and nuclear translocation. These data also demonstrate that our disease-specific biomarker identification approach through a comparison of multiple diseases is a powerful tool to identify unique pathomechanisms underlying the specific disease.

DISCUSSION

Databases for the dissemination of experimental data have been established by the bioinformatics community to facilitate data reuse and reanalysis, including the Gene Expression Omnibus by the National Center for Biotechnology Information and ArrayExpress by the European Bioinformatics Institute for experimental transcriptomics data. One of the biggest challenges that have limited the reuse of these data is that differences in expression data distributions due to technical differences in experimental methodologies make it difficult to perform true meta-analysis without the use of methods for cross-experiment data normalization. However, cross-experiment normalization approaches also run the risk of erasing true biological differences. In this study, we show that cross-comparison of different inflammatory skin disease transcriptomic studies using a unique bioinformatics normalization approach—control RatioA—identified disease-specific markers that were not obviously apparent in the evaluation of individual diseases in single experiments. In fact, many of the differentially expressed genes found in individual experiments reflect common inflammatory responses rather than disease-specific biomarkers. Importantly, this cross-experiment meta-analysis not only appears to provide biomarkers as a basis for a molecular skin disease classification but also reveals core pathogenic events, which can be further developed as therapeutic targets.

The control RatioA normalization approach used in this study differs from the RMA/frozen RMA approach used in two previously published meta-analyses of overlapping sets of skin diseases (Inkeles et al., 2015; Wong et al., 2012). RMA and frozen RMA are global normalization approaches that are designed to produce similar expression distributions across all samples, regardless of whether the sample is lesional or control, which is a reasonable approach if most genes are unaffected by the disease state. However, for some diseases evaluated in our studies, the number of genes that showed DE before normalization was relatively large (>22,000 for AD and >17,000 for psoriasis [Figure 1]), and so global normalization by RMA/frozen RMA runs the risk of normalizing away true disease-specific biological signal. By deriving gene-specific normalization factors on control samples alone, the control RatioA approach used in this study avoids this problem. Indeed, we show that the set of disease-specific, differentially expressed genes is largely retained after control RatioA normalization, which was not evaluated in the previous studies using RMA/frozen RMA normalization.

In dermatology, some 3,000 varieties of skin diseases have been described (Bickers et al., 2006), mostly defined by their morphological features and supported by histologic analysis that recognizes a limited number of anatomical patterns. The possibility to analyze 20,000 genes with the unbiased identification of molecular transcriptional abnormalities that define the disease would have tremendous diagnostic value in clinical practice as targets for immunohistochemistry evaluation of tissue biopsies or serum proteomics analysis. Indeed, the NS-Forest analysis identified minimum sets of marker genes that produce optimal classification accuracy that could be used for various diagnostic applications, which were validated in separate transcriptomics experiments using a different assay technology (RNA sequencing) and in protein-targeting immunohistochemical experiments.

In addition to the minimum set of diagnostic biomarkers whose combined expression pattern provides optimal classification, NS-Forest also produces an extended set of disease-specific, differentially expressed binary genes. For example, for psoriasis, the minimum marker gene classifier includes *GDA* and *LUZP2*, whose combined expression provides a classification accuracy >0.99. The association of these marker genes with psoriasis has not been previously reported, and little is known about their function in the pathogenesis of psoriasis. The extended NS-Forest binary gene list for psoriasis also includes *NOS2* (Quaranta et al., 2014), *PLA2G4D* (Shao et al., 2021), and *VNN3* (Jansen et al., 2009), which are well-known markers of psoriasis validating the overall analysis approach.

Although the use of NS-Forest for marker gene selection has been extensively validated and compared with other alternative approaches on multiple datasets (Aevermann et al., 2021), other methods may still be useful for specific use cases. Therefore, we also provide complete lists of differentially expressed genes specific to each of the inflammatory skin diseases (Supplementary Table S2), providing a more comprehensive view of the specific and distinct molecular pathways of each of these distinct inflammatory diseases.

The work described in this study is complementary to a recent study reporting the transcriptomic analysis of $CD45^+$ inflammatory cells from AD and psoriasis skin samples using single-cell RNA sequencing (Liu et al., 2022). Integration of these single-cell expressions into our bulk transcriptomics profiles could allow us to connect the expression of disease-specific markers with specific cell types either from the inflammatory infiltrate or from the resident skin cells. A preliminary comparison between the NS-Forest–extended binary marker gene list and the differentially expressed genes reported for the $CD45^+$ inflammatory cells in psoriasis (Swindell et al., 2013) shows no overlap (data not shown), suggesting that the predominant gene expression signal found in skin biopsy samples is derived from resident skin cells rather than from inflammatory infiltrate.

Our study provides evidence that the comparative multi-disease approach is a powerful tool to unravel disease pathomechanisms. In fact, we found that the specific CLE classifier PTEN is also a key molecule in the pathogenesis of CLE by its ability to promote type I IFNs. *PTEN* is a tumor suppressor gene that dephosphorylates PIP3 and thereby inhibits PI3K–protein kinase B (Akt) signaling. Loss of PTEN leads to uncontrolled PI3K–Akt signaling, with cell survival and cell growth in cancer (Lee et al., 2018), but the consequences of increased PTEN activity, as we observed in CLE, have not been described. In the context of viral infections, PTEN promotes type I IFN production in monocytes independently of PI3K/Akt but through dephosphorylation of IRF3 at Ser97, resulting in the release of IRF3 for nuclear import and activation (Li et al., 2016; Xu et al., 2021). Our study now identifies a similar PTEN-dependent IRF3 nuclear translocation in KCs and shows a link between the increased PTEN expression in CLE KCs and the aberrant IRF3 activation with pathogenic IFN-β and IFN-κ production.

Type I IFNs are key cytokines in the pathogenesis of CLE. Blocking type I IFN signaling with the anti-IFNAR antibody anifrolumab (Morand et al., 2020) has demonstrated efficacy in the treatment of CLE (Blum et al., 2022; Trentin et al., 2023). The principal sources of IFN-α in lupus are pDCs, specialized nucleic acid sensing through toll-like receptor 7/9 that infiltrate the dermis of CLE lesions (Farkas et al., 2001; Lande et al., 2011). Treatment of patients with CLE with the pDC-specific anti-BDCA2 antibody litifilimab has indeed shown therapeutic efficacy (Werth et al., 2022). However, several studies also point to KCs as an important source of pathogenic type I IFNs in CLE. KCs in CLE may be activated by double-stranded RNA products released during UV injury (a major driver of CLE) through the double-stranded RNA sensors toll-like receptor 3 or RIG-I/MAVS and produce IFN-β and the KC-specific IFN-κ (Sarkar et al., 2018; Scholtissek et al., 2017). It has been postulated that IFN-κ plays a critical role in the development of CLE lesions by enhancing UV-mediated KC apoptosis and increasing their responsiveness to pDC-derived IFN-α (Sarkar et al., 2018). Because increased PTEN expression drives cell apoptosis by antagonizing the PI3K–Akt signaling, it is possible that PTEN also participates directly in driving KC death, which may then release self-nucleic acids that trigger the subsequent activation of pDC.

The mechanisms underlying increased PTEN expression by CLE KCs are unknown. Interestingly, during viral infection, a long noncoding RNA induced by nucleic acid sensing or exposure to type I IFNs enhances PTEN expression and promotes IRF3 activation by releasing the inhibitory effect of hsa-miR-107 (Xu et al., 2021). Because an increased expression of a long noncoding RNA regulating PTEN in PBMCs has been linked to the risk of developing systemic LE (Liu et al., 2021), it is possible that similar mechanisms for PTEN upregulation occur in CLE KCs.

Our study warrants the evaluation of PTEN inhibitors for the treatment of CLE. Because both PI3K/Akt inhibition as well as IRF3 translocation are mediated by the phosphatase activity of PTEN, inhibitors need to target this domain. PTEN's risky profile as a tumor suppressor has to be considered. In this sense, potent inhibitors that allow selective, short-term PTEN inhibition preferably in a tissue-specific manner may be an optimal therapeutic approach for CLE.

In summary, we show that comparison of transcriptional profiles of multiple diseases from independent studies is a powerful tool to generate a molecular disease classification and unravel pathomechanism and potential therapeutic targets through the identification of disease-specific markers. All of the primary and derived data reported in this study have been made available through a public resource—the Skin Science Foundation Bioinformatics Hub—for use by the broader skin science community (https://biohub.skinsciencefoundation.org).

MATERIALS AND METHODS

Human samples and data sets

Studies were approved by the institutional review boards and the local ethics committee of the Lausanne University Hospital (CHUV) (Lausanne, Switzerland), in accordance with the Helsinki Declaration and were reviewed by the ethical committee board of the canton of Vaud, Switzerland (CER-VD 2020-02204). Biobanked formalin-fixed, paraffin-embedded skin tissues stored in the Swiss Biobanking Platform–accredited dermatology biobank were obtained from patients who provided written, informed consent for research use.

Landscape analysis

To begin building a knowledge resource about inflammatory skin diseases for the skin science community, we conducted a cross-comparison landscape meta-analysis of publicly available human-derived omics datasets. Publicly available candidate transcriptomic datasets were identified by searching Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) and ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) for human inflammatory skin diseases of interest, such as psoriasis, AD, LP, CLE, rosacea, vitiligo, AA, HS, and acne vulgaris (acne). In total, 200 candidate datasets were found in Gene Expression Omnibus and ArrayExpress, with the majority (128 datasets) profiling either psoriasis or AD. The remaining 72 profiled a mixture of the other diseases of interest.

For each disease of interest, one representative dataset was selected on the basis of the following criteria: samples were sourced from skin biopsies, healthy controls with or without nonlesional samples were included in addition to the lesional samples, and the dataset was associated with at least one peer-reviewed publication. Given that many disease types did not have publicly available RNA-sequencing data at the time, all datasets selected were generated using a microarray platform. In cases where more than one dataset

met the above criteria, the dataset with the most samples was selected. As a result, eight datasets representing nine skin diseases interrogating ~47,000 transcripts were selected for the cross-comparison landscape meta-analysis (Supplementary Table S3) and can be found at the Skin Science Bioinformatics Hub: https://biohub.skinsciencefoundation.org.

Data normalization for cross-experiment integration using control RatioA

Given that the experiments were conducted on different platforms by different research groups, the range of expression values was different in each of the different experiments (Figure 1a and b). Thus, to integrate all eight experiments selected for meta-analysis, the expression data were normalized using a modification of the RatioA approach under the assumption that the expression level of any given gene should be the same in the control samples across all experiments, that is, control RatioA:

$$Y_i = \log_2(X_i / \text{mean}[X_{i\ \text{control}}])$$

For each gene (Y_i), the mean expression across all healthy control samples was calculated (mean $[X_{i\ \textbf{control}}]$) for each experiment independently. This experiment-specific mean expression of control samples was then used as an experiment-specific normalization factor by dividing it into every expression value (X_i) for that gene in every sample, including healthy, lesional, and nonlesional, for that experiment. Log_2 transformation of this ratio results in a relatively symmetric distribution of normalized expression values for control samples centered around 0 in every experiment (Supplementary Figure S2) while retaining better DE in test samples (Figure 1c) in comparison with other normalization methods.

To determine whether the normalization approach had any effect on the underlying biological relationships captured in the experiment, DE analysis was performed both before (pre) and after (post) normalization for each experiment using a standard pairwise DE analysis (Limma) (Ritchie et al., 2015) comparing the disease and control samples (Figure 1c).

Transcriptome-based sample clustering

Unsupervised clustering was performed on the control RatioA-normalized gene expression matrix using the R package SC3 (Kiselev et al., 2017) (Figure 2a). SC3 uses a consensus clustering approach, in which many different parameters and methods are explored, and a final solution is determined from the consensus result. In this study, we used the default consensus parameters while generating solutions for 3−15 (k) clusters. The value of k was determined as the maximum average silhouette score for each of the clusters (as described in Kiselev et al., 2017), a measure of separation and distinction, ranging from −1 to 1, with the highest score of 0.72 obtained with $k = 9$. In addition to k-means clustering, SC3 also hierarchically clusters each k-means solution. The hierarchical dendrogram corresponding to the best k-means solution was then considered as capturing the taxonomic relationships among skin diseases on the basis of their transcriptional profiles.

Disease-specific gene expression signature determination

DE analysis between diseases was performed on the resulting SC3 taxonomy using Limma. First, all healthy control samples were removed from the taxonomy because we were primarily interested in finding gene signatures that distinguished between the different disease samples rather than between disease samples and healthy controls. Next, DE was computed by comparing each node of interest in the taxonomy with all other samples that were not in that node (a one-versus-all approach). DE results were filtered by positive-only fold change and an adjusted $P < 0.05$ (Bonferroni correction). Positive fold-change results were chosen because when comparing a disease with other diseases, the genes with negative fold-change values were found to be the positively associated genes for the other diseases.

Pathway analysis was conducted using the DE findings, gene ontology annotations, and the DAVID (Database for Annotation, Visualization and Integrated Discovery) resource (Huang da et al., 2009a, 2009b) (https://david.abcc.ncifcrf.gov/); additional pathway enrichment analysis was performed using the Reactome resource (Jassal et al., 2020) (https://reactome.org/).

Marker gene determination

Marker gene combinations were identified using NS-Forest, version 2.0 (Aevermann et al., 2021) (https://github.com/JCVenterInstitute/NSForest). NS-Forest uses random forest classification model construction and feature selection to determine two sets of markers of a given class—in this case, disease type with (i) the minimum set of most discriminatory gene combinations and (ii) an extended set of binary markers available for a given class. NS-Forest has been extensively validated using real and simulated datasets and favorably compares with alternative marker selection methods (Aevermann et al., 2021). NS-Forest was run using default parameters on the control RatioA-normalized gene expression data.

KC stimulation

Normal human epidermal KCs (Sigma-Aldrich, St. Louis, MO) cultured in KC serum-free medium containing antibiotics (1% penicillin/streptomycin) supplemented with human recombinant epidermal GF and bovine pituitary extract (Gibco, Thermo Fisher Scientific, Waltham, MA) were used. KCs were pretreated or not with the PTEN inhibitor bpV(pic) (5 μM, Sigma-Aldrich) for 1 hour, followed by stimulation with liposomes containing poly(I:C) (2 μg/ml, InvivoGen, San Diego, CA). Cells were harvested 4 hours later for immunofluorescence staining or 24 hours later for RT-qPCR analysis. For RNA-silencing experiments, HaCaT cells cultured in DMEM 10% fetal bovine serum were treated with liposomes containing 1 mg of siRNA control (SignalSilence Control siRNA #6568, Cell Signaling Technology, Danvers, MA) or siRNA targeting PTEN (SignalSilence PTEN siRNA I #6251, Cell Signaling Technology) for 2 days. PTEN expression silencing was assessed by RT-qPCR and immunofluorescence staining, and siRNA-treated cells were further stimulated as discussed earlier.

Immunohistochemistry analysis

Formalin-fixed, paraffin-embedded skin blocks from different patients with inflammatory skin diseases were cut into 6-μm sections and placed on slides. Sections were first deparaffinized and rehydrated, and then heat-induced epitope retrieval was performed, and sections were permeabilized with PBS 0.01% Triton. Samples were stained with rabbit anti-human PTEN (recombinant rabbit monoclonal #EPR9941-2, Abcam, Cambridge, United Kingdom; 1/50) for 2 hours at room temperature. Sections were then stained with ImmPRESS horse radish peroxide horse anti-rabbit IgG (ready to use, Vector Laboratories, Newark, CA) followed by 3,3′-diaminobenzidine staining and Mayer counterstaining. Slides were digitalized using the PANNORAMIC 250 Flash digital scanner (3DHISTECH, Budapest, Hungary), and staining intensity was quantified using ImageJ software (National Institutes of Health, Bethesda, MD).

Immunofluorescence analysis

Normal human epidermal KCs were cultured onto poly-L-lysine–pretreated Teflon slides and stimulated as discussed earlier. Cells were then fixed with Intracellular Fixation Buffer and stained with antiphosphorylated IRF3 (1/200, Cell Signaling Technology) for 1 hour at room temperature. Cells were then stained with A488-labeled mouse anti-rabbit IgG antibodies (1/500, Thermo Fisher Scientific) for 30 minutes at room temperature. For immunofluorescence analysis of CLE skin sections, formalin-fixed, paraffin-embedded skin blocks were prepared as discussed earlier and stained with mouse anti-human PTEN (clone 1B8, 1/100, Invitrogen, Waltham, MA) and rabbit anti-human IFN-β (1/1,000, Invitrogen) followed by A488-labeled donkey anti-mouse IgG (1/500) and A546-labeled donkey anti-rabbit IgG (1/500) antibodies (Thermo Fisher Scientific) or with rabbit anti-human PTEN (recombinant rabbit monoclonal #EPR9941-2, 1/50, Abcam) and mouse anti-human IFN-κ (monoclonal mouse IgG2A clone #1009725, 1/100, R&D Systems, Minneapolis, MN) followed by A488-labeled donkey anti-rabbit IgG (1/500) and A546-labeled donkey anti-mouse IgG (1/500) antibodies (Thermo Fisher Scientific). Images of the different stained cells or skin sections were acquired with a Zeiss LSM 700 confocal microscope and analyzed with the Fiji software. Multicolor images were first split into the different fluorescence channels to generate images of each marker, and selection areas were drawn for each individual cell (or each individual nucleus) using the multicolor image and saved to the region of interest manager to further apply these regions of interests to the fluorescence channel to be quantified. Fluorescence intensities were then measured for each region of interest, and correlation analysis between markers was performed in GraphPad Prism 9 (GraphPad Software, San Diego, CA).

Data availability statement

All gene expression and differential expression data are made freely available for exploration at the Skin Science Foundation Bioinformatic Hub (https://biohub.skinsciencefoundation.org). All the accession numbers of datasets used in this study are given in Supplementary Table S3.

CONFLICT OF INTEREST

HB reports paid consultancies for Anaptysbio, Boehringer Ingelheim, Bristol Myers Squibb, Eli-Lilly, Janssen, and UCB and grant support by Boehringer Ingelheim, Bristol Myers Squibb, Janssen, and Pfizer. JKK is an employee of Galderma. TL is employed and funded by LEO Pharma. CEMG has received honoraria and/or research grants from Almirall, Amgen, BMS, Boehringer-Ingelheim, Dermavant, Eli Lilly, GSK, Janssen, Novartis, ONO Pharmaceutical, and UCB. The remaining authors state no conflict of interest.

ACKNOWLEDGMENTS

This work was performed on behalf of the Skin Science Foundation. This work was supported by a grant from the Skin Science Foundation to BDA and RHS.

AUTHOR CONTRIBUTIONS

Conceptualization: RHS, MG; Data Curation: BDA, JDD, FS; Supervision: RHS, MG; Writing – Original Draft Preparation: RHS, MG; Writing – Review and Editing: JMA, HB, JB, MH, VJ, KJ, JKK, TL, IP, KYS, MSc, MSie, MSim

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at https://doi.org/10.1016/j.jid.2023.06.211.

REFERENCES

Aevermann B, Zhang Y, Novotny M, Keshk M, Bakken T, Miller J, et al. A machine learning method for the discovery of minimum marker gene combinations for cell type identification from single-cell RNA sequencing. Genome Res 2021;31:1767–80.

Banchereau R, Cepika AM, Banchereau J, Pascual V. Understanding human autoimmunity and autoinflammation through transcriptomics. Annu Rev Immunol 2017;35:337–70.

Banchereau R, Hong S, Cantarel B, Baldwin N, Baisch J, Edens M, et al. Personalized immunomonitoring uncovers molecular networks that stratify lupus patients. Cell 2016;165:1548–50.

Baurecht H, Hotze M, Brand S, Büning C, Cormican P, Corvin A, et al. Genome-wide comparative analysis of atopic dermatitis and psoriasis gives insight into opposing genetic mechanisms [published correction appears in Am J Hum Genet 2015;97:933] Am J Hum Genet 2015;96:104–20.

Bickers DR, Lim HW, Margolis D, Weinstock MA, Goodman C, Faulkner E, et al. The burden of skin diseases: 2004 a joint project of the American Academy of Dermatology Association and the Society for Investigative Dermatology. J Am Acad Dermatol 2006;55:490–500.

Bieber T. Atopic dermatitis: an expanding therapeutic pipeline for a complex disease. Nat Rev Drug Discov 2022;21:21–40.

Blanco P, Palucka AK, Gill M, Pascual V, Banchereau J. Induction of dendritic cell differentiation by IFN-alpha in systemic lupus erythematosus. Science 2001;294:1540–3.

Blok JL, Li K, Brodmerkel C, Jonkman MF, Horváth B. Gene expression profiling of skin and blood in hidradenitis suppurativa. Br J Dermatol 2016;174:1392–4.

Blum FR, Sampath AJ, Foulke GT. Anifrolumab for treatment of refractory cutaneous lupus erythematosus. Clin Exp Dermatol 2022;47:1998–2001.

Buhl T, Sulk M, Nowak P, Buddenkotte J, McDonald I, Aubert J, et al. Molecular and morphological characterization of inflammatory infiltrate in rosacea reveals activation of Th1/Th17 pathways. J Invest Dermatol 2015;135:2198–208.

Chasset F, Francès C. Current concepts and future approaches in the treatment of cutaneous lupus erythematosus: a comprehensive review. Drugs 2019;79:1199–215.

Conrad C, Gilliet M. Psoriasis: from pathogenesis to targeted therapies. Clin Rev Allergy Immunol 2018;54:102–13.

De Marco C, Laudanna C, Rinaldo N, Oliveira DM, Ravo M, Weisz A, et al. Specific gene expression signatures induced by the multiple oncogenic alterations that occur within the PTEN/PI3K/AKT pathway in lung cancer. PLoS One 2017;12:e0178865.

Eyerich K, Eyerich S. Immune response patterns in non-communicable inflammatory skin diseases. J Eur Acad Dermatol Venereol 2018;32:692–703.

Farkas L, Beiske K, Lund-Johansen F, Brandtzaeg P, Jahnsen FL. Plasmacytoid dendritic cells (natural interferon- alpha/beta-producing cells) accumulate in cutaneous lupus erythematosus lesions. Am J Pathol 2001;159:237–43.

Frisoli ML, Essien K, Harris JE. Vitiligo: mechanisms of pathogenesis and treatment. Annu Rev Immunol 2020;38:621–48.

Fyhrquist N, Muirhead G, Prast-Nielsen S, Jeanmougin M, Olah P, Skoog T, et al. Microbe-host interplay in atopic dermatitis and psoriasis. Nat Commun 2019;10:4703.

Garcia-Romo GS, Caielli S, Vega B, Connolly J, Allantaz F, Xu Z, et al. Netting neutrophils are major inducers of type I IFN production in pediatric systemic lupus erythematosus. Sci Transl Med 2011;3:73ra20.

Goldburg SR, Strober BE, Payette MJ. Hidradenitis suppurativa: current and emerging treatments. J Am Acad Dermatol 2020;82:1061–82.

Griffiths CEM, Armstrong AW, Gudjonsson JE, Barker JNWN. Psoriasis. Lancet 2021;397:1301–15.

Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 2009a;37:1–13.

Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using David bioinformatics resources. Nat Protoc 2009b;4:44–57.

Inkeles MS, Scumpia PO, Swindell WR, Lopez D, Teles RMB, Graeber TG, et al. Comparison of molecular signatures from multiple skin diseases identifies mechanisms of immunopathogenesis. J Invest Dermatol 2015;135:151–9.

Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 2003;31:e15.

Jabbari A, Cerise JE, Chen JC, Mackay-Wiggan J, Duvic M, Price V, et al. Molecular signatures define alopecia areata subtypes and transcriptional biomarkers. EBiomedicine 2016;7:240–7.

Jansen PA, Kamsteeg M, Rodijk-Olthuis D, van Vlijmen-Willems IM, de Jongh GJ, Bergers M, et al. Expression of the vanin gene family in normal and inflamed human skin: induction by proinflammatory cytokines. J Invest Dermatol 2009;129:2167–74.

Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The reactome pathway KnowledgeBase. Nucleic Acids Res 2020;48: D498–503.

Kelhälä HL, Palatsi R, Fyhrquist N, Lehtimäki S, Väyrynen JP, Kallioinen M, et al. IL-17/Th17 pathway is activated in acne lesions. PLoS One 2014;9: e105238.

Kiselev VY, Kirschner K, Schaub MT, Andrews T, Yiu A, Chandra T, et al. SC3: consensus clustering of single-cell RNA-seq data. Nat Methods 2017;14: 483–6.

Lande R, Ganguly D, Facchinetti V, Frasca L, Conrad C, Gregorio J, et al. Neutrophils activate plasmacytoid dendritic cells by releasing self-DNA-peptide complexes in systemic lupus erythematosus. Sci Transl Med 2011;3:73ra19.

Lee YR, Chen M, Pandolfi PP. The functions and regulation of the PTEN tumour suppressor: new modes and prospects. Nat Rev Mol Cell Biol 2018;19:547–62.

Li S, Zhu M, Pan R, Fang T, Cao YY, Chen S, et al. The tumor suppressor PTEN has a critical role in antiviral innate immunity. Nat Immunol 2016;17: 241–9.

Liu CH, Lu YL, Huang HT, Wang CF, Luo HC, Wei GJ, et al. Association of LncRNA-GAS5 gene polymorphisms and PBMC LncRNA-GAS5 level with risk of systemic lupus erythematosus in Chinese population. J Cell Mol Med 2021;25:3548–59.

Liu J, Berthier CC, Kahlenberg JM. Enhanced inflammasome activity in systemic lupus erythematosus is mediated via Type I interferon-induced up-regulation of interferon regulatory Factor 1. Arthritis Rheumatol 2017;69: 1840–9.

Liu Y, Wang H, Taylor M, Cook C, Martínez-Berdeja A, North JP, et al. Classification of human chronic inflammatory skin disease based on single-cell immune profiling. Sci Immunol 2022;7:eabl9165.

Min X, Zheng M, Yu Y, Wu J, Kuang Q, Hu Z, et al. Ultraviolet light induces HERV expression to activate RIG-I signalling pathway in keratinocytes. Exp Dermatol 2022;31:1165–76.

Mo A, Marigorta UM, Arafat D, Chan LHK, Ponder L, Jang SR, et al. Disease-specific regulation of gene expression in a comparative analysis of juvenile idiopathic arthritis and inflammatory bowel disease. Genome Med 2018;10:48.

Morand EF, Furie R, Tanaka Y, Bruce IN, Askanase AD, Richez C, et al. Trial of anifrolumab in active systemic lupus erythematosus. N Engl J Med 2020;382:211–21.

Mylonas A, Hawerkamp HC, Wang Y, Chen J, Messina F, Demaria O, et al. Type I IFNs link skin-associated dysbiotic commensal bacteria to pathogenic inflammation and angiogenesis in rosacea. JCI Insight 2023;8: e151846.

Nestle FO, Conrad C, Tun-Kyi A, Homey B, Gombert M, Boyman O, et al. Plasmacytoid predendritic cells initiate psoriasis through interferon-alpha production. J Exp Med 2005;202:135–43.

Puar N, Chovatiya R, Paller AS. New treatments in atopic dermatitis. Ann Allergy Asthma Immunol 2021;126:21–31.

Quaranta M, Knapp B, Garzorz N, Mattii M, Pullabhatla V, Pennino D, et al. Intraindividual genome expression analysis reveals a specific molecular signature of psoriasis and eczema. Sci Transl Med 2014;6:244ra90.

Regazzetti C, Joly F, Marty C, Rivier M, Mehul B, Reiniche P, et al. Transcriptional analysis of vitiligo skin reveals the alteration of WNT pathway: a promising target for repigmenting vitiligo patients. J Invest Dermatol 2015;135:3105–14.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. LIMMA powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.

Rönnblom L, Eloranta ML, Alm GV. Role of natural interferon-alpha producing cells (plasmacytoid dendritic cells) in autoimmunity. Autoimmunity 2003;36:463–72.

Sarkar MK, Hile GA, Tsoi LC, Xing X, Liu J, Liang Y, et al. Photosensitivity and type I IFN responses in cutaneous lupus are driven by epidermal-derived interferon kappa. Ann Rheum Dis 2018;77:1653–64.

Schmid AC, Byrne RD, Vilar R, Woscholski R. Bisperoxovanadium compounds are potent PTEN inhibitors. FEBS Lett 2004;566:35–8.

Scholtissek B, Zahn S, Maier J, Klaeschen S, Braegelmann C, Hoelzel M, et al. Immunostimulatory endogenous nucleic acids drive the lesional inflammation in cutaneous lupus erythematosus. J Invest Dermatol 2017;137: 1484–92.

Shao S, Chen J, Swindell WR, Tsoi LC, Xing X, Ma F, et al. Phospholipase A2 enzymes represent a shared pathogenic pathway in psoriasis and Pityriasis rubra pilaris. JCI Insight 2021;6:e151911.

Shao S, Tsoi LC, Sarkar MK, Xing X, Xue K, Uppala R, et al. IFN-γ enhances cell-mediated cytotoxicity against keratinocytes via JAK2/STAT1 in lichen planus. Sci Transl Med 2019;11:eaav7561.

Swindell WR, Johnston A, Voorhees JJ, Elder JT, Gudjonsson JE. Dissecting the psoriasis transcriptome: inflammatory- and cytokine-driven gene expression in lesions from 163 patients. BMC Genomics 2013;14:527.

Trentin F, Tani C, Elefante E, Stagnaro C, Zucchi D, Mosca M. Treatment with anifrolumab for discoid lupus erythematosus. JAMA Dermatol 2023;159: 224–6.

Vallin H, Blomberg S, Alm GV, Cederblad B, Rönnblom L. Patients with systemic lupus erythematosus (SLE) have a circulating inducer of interferon-alpha (IFN-alpha) production acting on leucocytes resembling immature dendritic cells. Clin Exp Immunol 1999;115:196–202.

Wang EHC, Sallee BN, Tejeda CI, Christiano AM. JAK inhibitors for treatment of alopecia areata. J Invest Dermatol 2018;138:1911–6.

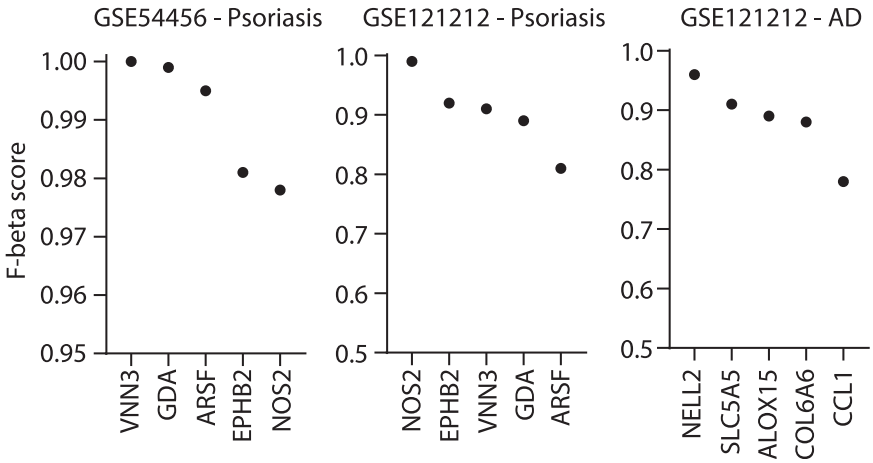
Wang L, Oh WK, Zhu J. Disease-specific classification using deconvoluted whole blood gene expression. Sci Rep 2016;6:32976.

Wenzel J. Cutaneous lupus erythematosus: new insights into pathogenesis and therapeutic strategies. Nat Rev Rheumatol 2019;15:519–32.

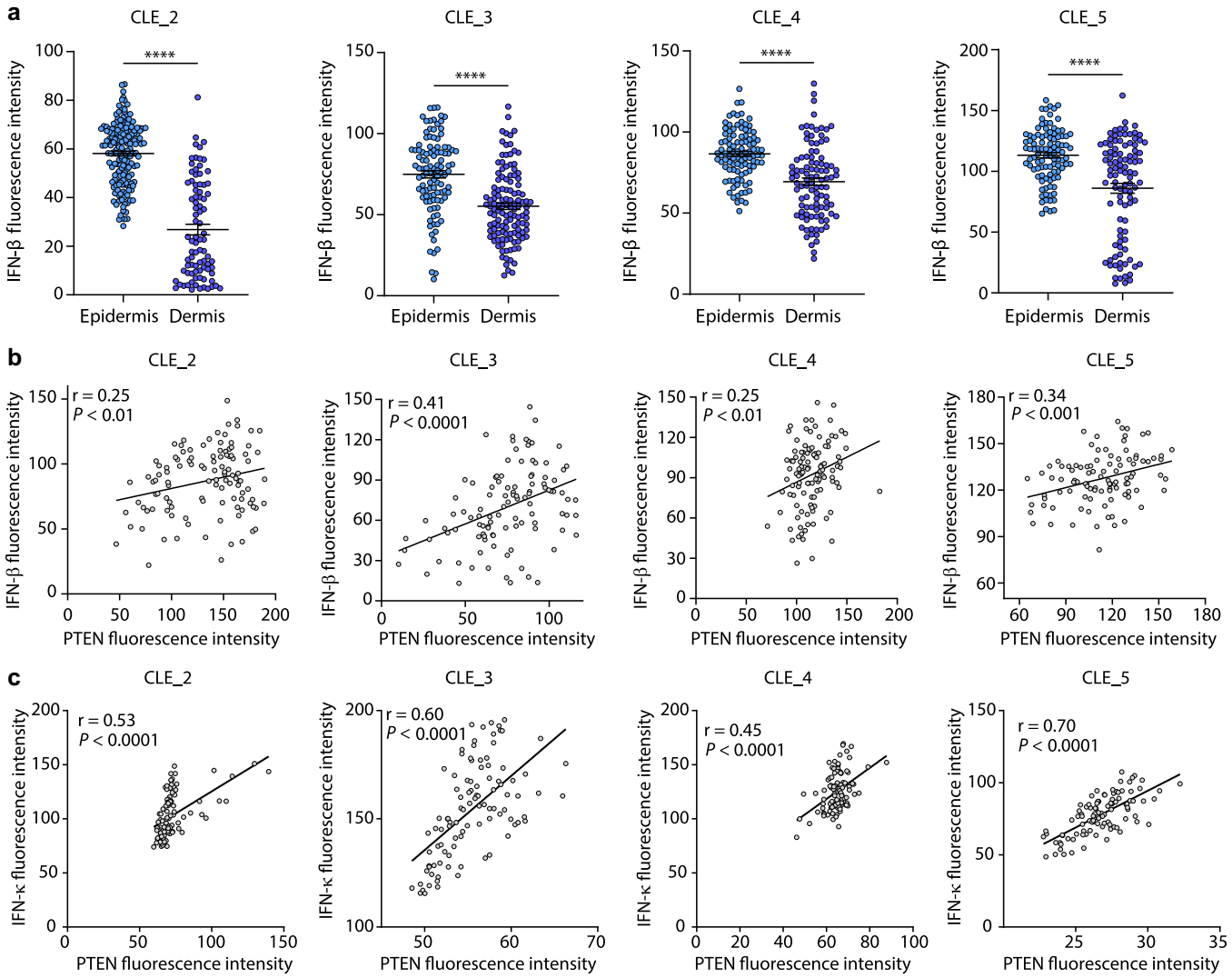
Werth VP, Barbey C, Franchimont N. Anti-BDCA2 antibody for cutaneous lupus erythematosus. Reply. N Engl J Med 2022;387:1529.

Wong D, Kea B, Pesich R, Higgs BW, Zhu W, Brown P, et al. Interferon and biologic signatures in dermatomyositis skin: specificity and heterogeneity across diseases. PLoS One 2012;7:e29161.

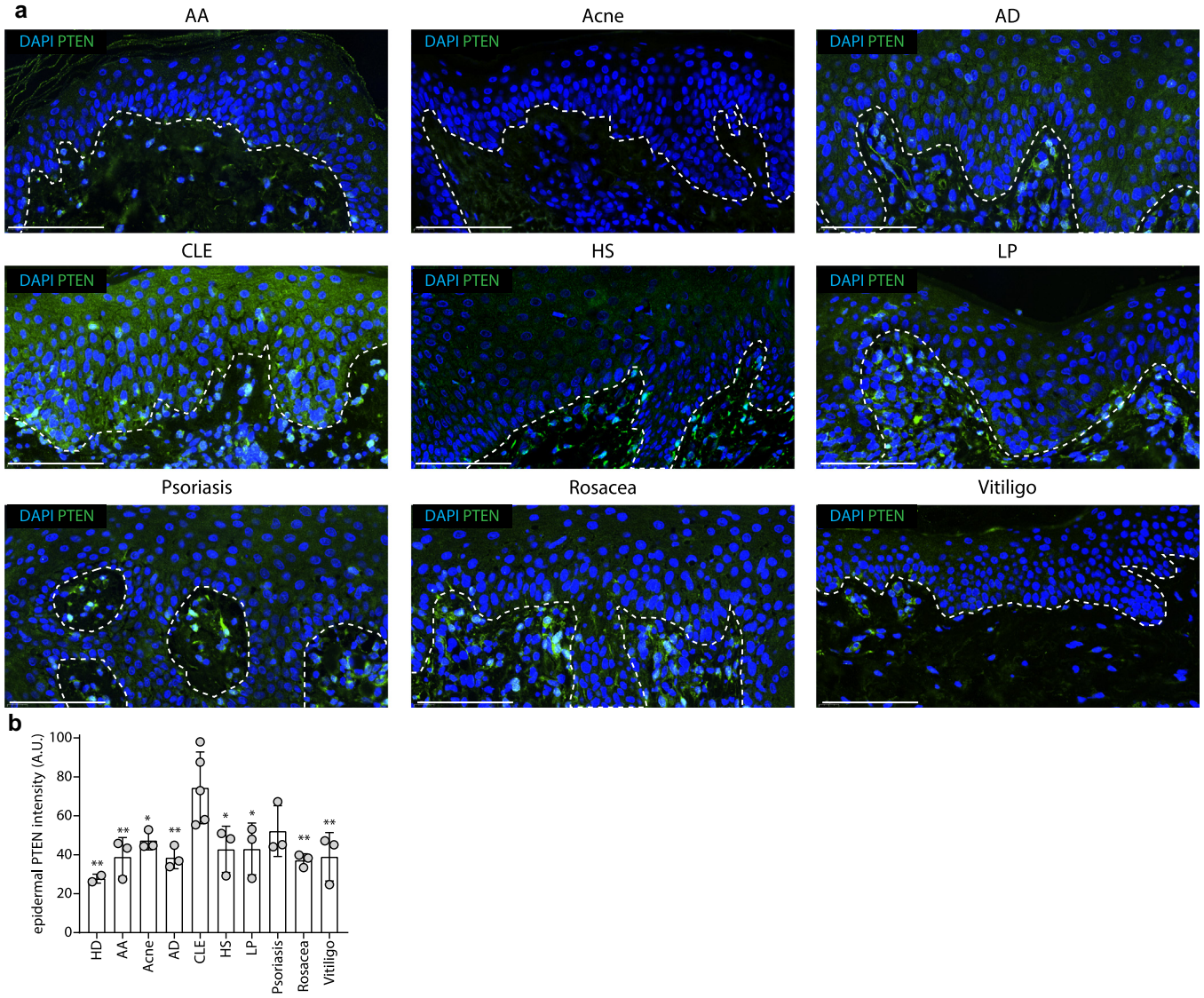
Xu R, Yu SS, Yao RR, Tang RC, Liang JW, Pang X, et al. Interferon-inducible LINC02605 promotes antiviral innate responses by strengthening IRF3 nuclear translocation. Front Immunol 2021;12:755512.



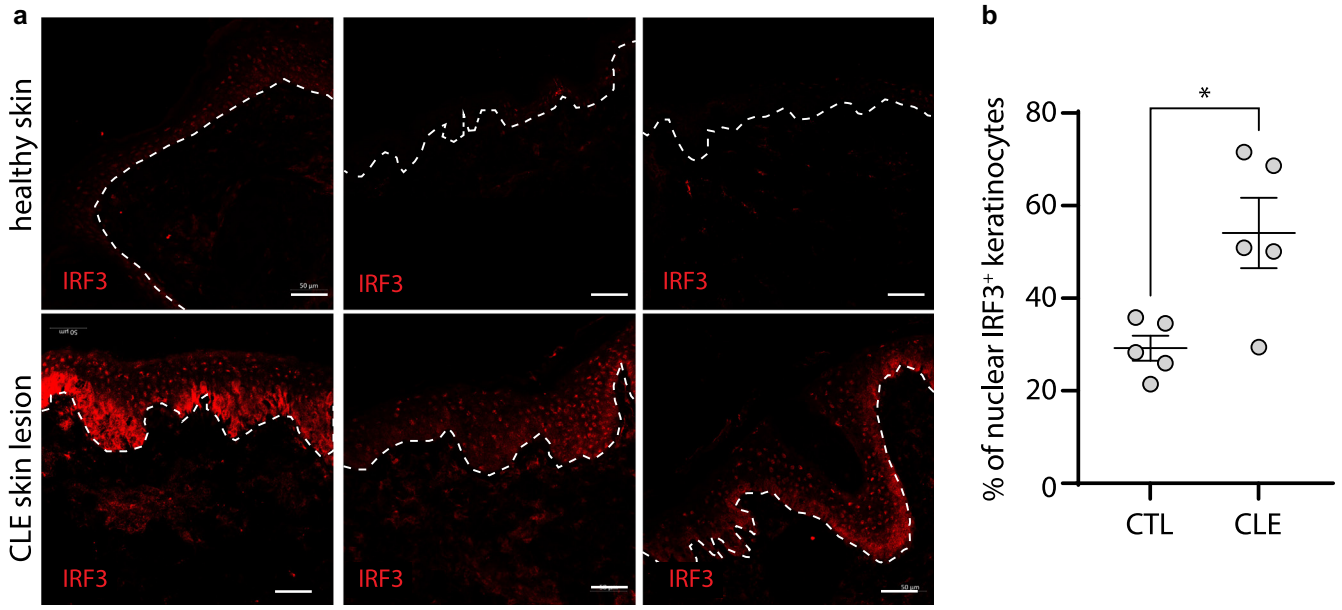
Supplementary Figure S1. Validation of NS-Forest binary markers with RNA-seq datasets. Validation of the PSO- and AD-associated NS-Forest binary markers using the independent GSE54456 and GSE121212 bulk RNA-seq datasets is shown. Classification accuracy using expression of single marker genes between lesional and healthy control samples as quantified by F-beta score is shown. AD, atopic dermatitis; PSO, psoriasis; RNA-seq, RNA sequencing.



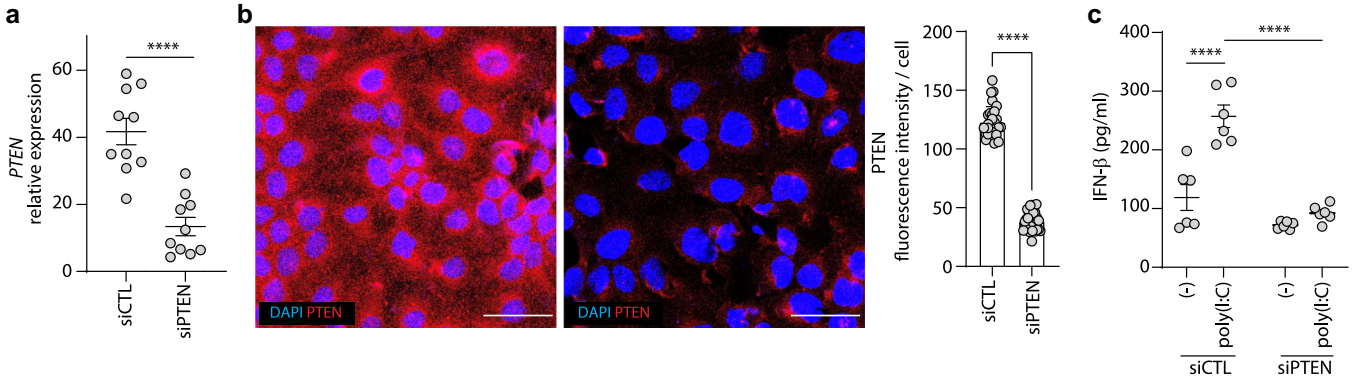
Supplementary Figure S2. PTEN expression correlates with type I IFN production by keratinocytes in CLE. (**a**) Quantification of IFN-β fluorescence intensity at the single-cell level in epidermal and dermal cells of four different lesional CLE skin specimens. Bars represent the mean ± SEM. Data were statistically analyzed using unpaired *t*-test. ****$P < 0.0001$. (**b**, **c**) Correlation between PTEN and (**b**) IFN-β or (**c**) IFN-κ fluorescence intensities of four different lesional CLE skin specimens at the single-cell level is shown. Pearson correlation coefficient and a two-tailed statistical significance are given. AA, alopecia areata; AD, atopic dermatitis; AU, arbitrary unit; CLE, cutaneous lupus erythematosus; HD, healthy donor; HS, hidradenitis suppurativa; LP, lichen planus.



Supplementary Figure S3. Increased expression of PTEN in keratinocytes of CLE. (**a**) Immunofluorescence staining of PTEN in skin sections from the different patients and (**b**) quantification of its intensity in the epidermis. Bars = 100 mm. Data were statistically analyzed using one-way ANOVA followed by Dunnett's multiple comparisons test. CLE was used as a reference. $^{**}P < 0.005$ and $^{*}P < 0.05$. CLE, n = 5; other diseases, n = 3. CLE, cutaneous lupus erythematosus.



Supplementary Figure S4. IRF3 is present in the nuclei of keratinocytes in CLE skin lesions. (**a**) Representative immunofluorescence stainings of IRF3 in skin sections from healthy donors (top, n = 5) and patients with CLE (bottom, n = 5). The dashed line delineates the dermo–epidermal junction. Bars = 50 mm. (**b**) Percentages of keratinocytes with nuclear IRF3 in healthy (CTL) and CLE skin from images in **a**. Data were statistically analyzed using two-tailed unpaired *t*-test. $*P < 0.05$. CLE, cutaneous lupus erythematosus; CTL, control.



Supplementary Figure S5. PTEN controls type I IFN production in keratinocytes. (**a**) *PTEN* mRNA expression in HaCaT cells treated with siPTEN or siCTL. (**b**) Immunofluorescence staining of PTEN in HaCaT cells treated as in **a** (left) and quantification of the fluorescence intensity at the single-cell level (right). Bars = 50 mm. (**c**) Production of IFN-β by HaCaT cells pretreated with siRNAs and stimulated with empty liposomes (−) or liposomes containing poly(I:C). (**a, b**) Data were statistically analyzed using unpaired *t*-test. $****P < 0.0001$. (**c**) Data were statistically analyzed using a two-way ANOVA followed by Sidak's multiple comparisons test. siCTL, control-targeted small interfering RNA; siPTEN, PTEN-targeted small interfering RNA; siRNA, small interfering RNA.