White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/154208/

Version: Published Version

Article:
Reimann, E., Lättekivi, F., Keermann, M. et al. (2019) Multicomponent biomarker approach improves the accuracy of diagnostic biomarkers for psoriasis vulgaris. Acta Dermato Venereologica, 99 (13). pp. 1258-1265. ISSN: 0001-5555

https://doi.org/10.2340/00015555-3337

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

INVESTIGATIVE REPORT

Check for updates

Multicomponent Biomarker Approach Improves the Accuracy of Diagnostic Biomarkers for Psoriasis Vulgaris

Ene REIMANN[1#], Freddy LÄTTEKIVI[1#], Maris KEERMANN[2,3], Kristi ABRAM[2,3], Sulev KÕKS[4,5], Külli KINGO[2,3] and Alireza FAZELI[1,6]
[1]Department of Pathophysiology, [2]Department of Dermatology, University of Tartu, [3]Clinic of Dermatology, Tartu University Hospital, Tartu, Estonia, [4]Centre for Comparative Genomics, Murdoch University, [5]The Perron Institute for Neurological and Translational Science, Perth, Australia and [6]Department of Oncology and Metabolism, The Medical School, University of Sheffield, Sheffield, UK
[#]Both authors contributed equally to this manuscript.

Accurate biomarker-based diagnosis of psoriasis vulgaris has remained a challenge; no reliable disease-specific biomarkers have yet been identified. There are several different chronic inflammatory skin diseases that can present similar clinical and dermoscopy features to psoriasis vulgaris, making accurate diagnosis more difficult. Both literature-based and data-driven selection of biomarker was conducted to select candidates for a multicomponent biomarker for psoriasis vulgaris. Support vector machine-based classification models were trained using gene expression data from locally recruited patients and validated on 7 public datasets, which included gene expression data of other inflammatory skin diseases in addition to psoriasis vulgaris. The resulting accuracy of the best classification model based on the expression levels of 4 genes (*IL36G, CCL27, NOS2* and *C10orf99*) was 96.4%, outperforming classification based on other marker gene combinations, which were more affected by variability in gene expression profiles between different datasets and patient groups. This approach has the potential to fill the void of clinically applicable diagnostic biomarkers for psoriasis vulgaris and other inflammatory skin diseases.

Key words: psoriasis; transcriptome; support vector machine.

Accepted Oct 14, 2019: E-published Oct 14, 2019

Acta Derm Venereol 2019; 99: 1258–1265.

Corr: Freddy Lättekivi, Department of Pathophysiology, University of Tartu, 14b Ravila Str., EE-50411 Tartu, Estonia. E-mail: freddy.lattekivi@ut.ee

SIGNIFICANCE

This article highlights the issue that previously proposed gene expression-based biomarkers for psoriasis vulgaris are dataset-specific, and therefore perform less accurately in alternative groups of patients. This paper addresses this problem by demonstrating that combining several potential biomarkers into a single multicomponent biomarker results in more accurate classification models, which retain their accuracy across different datasets and patient groups with the inclusion of other inflammatory skin diseases. We believe that this approach has the potential to fill the void of clinically applicable diagnostic biomarkers for psoriasis vulgaris and other inflammatory skin diseases.

Psoriasis can be a challenging disease to diagnose, as no reliable disease-specific and clinically viable biomarkers have been identified. Psoriasis is one of the most prevalent chronic inflammatory autoimmune diseases of the skin, affecting 2–3% of the population worldwide (1). Clinically, psoriasis can manifest itself in a broad spectrum of subtypes, the most common of which is psoriasis vulgaris (PsV) (2). PsV can share several similar clinical and molecular features with other chronic inflammatory skin conditions, such as parapsoriasis, lichen planus, pityriasis rosea, contact eczema, and atopic dermatitis (3–5). All these diseases, however, have different treatment and disease management strategies (6–9), which necessitates accurate and biomarker-based diagnosis.

Currently, the diagnosis of PsV relies mainly on the assessment of visible and dermoscopic symptoms by a clinician (10–13) or, in borderline cases, histological evaluation (2). Disease-specific biomarkers could also improve the diagnosis of PsV overlapping with other papulosquamous disorders and provide more accurate means of quantifying treatment efficacy. Years of high-throughput "omics" research have resulted in multiple potential biomarkers for PsV and other inflammatory skin diseases, yet most of them are universal for inflammation and no single marker has shown to be robustly specific to a given disease (14).

The approach of using multiple biomarkers in combination has proven to be more robust and accurate than the standalone measurements of individual biomarkers in the context of several human diseases (15–17). With the advancement of computational methods, multicomponent biomarkers offer a solution to overcome interindividual variability in the biomarker-based diagnostics of complex heterogeneous diseases (18–20). As such, we hypothesized that using a set of biomarkers in combination as a multicomponent biomarker would allow PsV to be distinguished from other inflammatory skin diseases more accurately than any of the individual markers independently. Coupling this with a low-cost gene expression quantification method, such as quantitative real-time PCR (qPCR) could result in a clinically viable tool for the diagnosis of PsV.

doi: 10.2340/00015555-3337
Acta Derm Venereol 2019; 99: 1258–1265

This is an open access article under the CC BY-NC license. www.medicaljournals.se/acta
Journal Compilation © 2019 Acta Dermato-Venereologica.

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

MATERIALS AND METHODS

Patients and sample collection

Twenty patients with psoriasis vulgaris (PsV) and 18 with atopic dermatitis (AD) were recruited from the Dermatology Clinic of Tartu University Hospital. The Ethics Review Committee on Human Research of the University of Tartu approved the protocols and informed consent forms used in this study. All participants signed an informed consent form. Detailed characteristics of recruited patients are shown in Appendix S1[1]. All skin samples were obtained with 4-mm punch biopsy. The biopsy specimens were instantly frozen on dry ice and stored at –80°C (for up to 3 months) until RNA extraction.

Public datasets included in the analysis

The NCBI GEO (https://www.ncbi.nlm.nih.gov/geo) datasets included in the analysis are summarized in **Table I**. The core datasets used for testing the disease-specific multicomponent biomarker approach were the microarray dataset GSE63741 and RNA-seq dataset GSE121212, as these included several other inflammatory skin diseases in addition to PsV. Additional 5 RNA-seq datasets were included in the test set in order to incorporate more of different PsV patient groups and additional skin conditions, such as atopic dermatitis (AD), non-lesional skin samples of AD patients (AD-NL) and non-lesional psoriatic skin (Ps-NL) samples. No datasets including other inflammatory skin diseases (e.g. lichen planus or contact eczema) were found among the RNA-seq datasets listed in the NCBI GEO repository.

Quantitative PCR analysis

Quantitative real-time PCR (qPCR) was used to quantify the expression levels of marker genes in 2 groups of locally recruited patients. Skin samples from the first group were collected as a part of a previous study (21) and consisted of 12 paired PsV and Ps-NL skin samples from the same patients and 12 samples from healthy controls (C). This dataset is hereinafter referred to as "qPCR dataset 1". The same samples also constitute the GSE66511 RNA-seq dataset. The second group consisted of 20 patients with PsV and 18 patients with AD. Gene expression data of these patients are hereinafter referred to as "qPCR dataset 2". The qPCR experiments were conducted following MIQE guidelines (22). Detailed description of the methods are shown in Appendix S1[1].

RNA-seq data analysis

All base-space datasets were aligned to the hg38 reference genome using STAR v2.6 (23). Aligned reads were quantified according to GENCODE (24) release 26 with the "–quantMode GeneCounts" option of STAR. Non-default options used for read alignment are shown in Appendix S1[1]. The colour-space SOLiD (Life Technologies Ltd) reads of the GSE66511 dataset were mapped to the hg19 reference genome and exon-exon junctions using LifeScope software (Life Technologies Ltd) mapping module with recommended settings. Mapping quality cut-off was set at 10 (MAPQ ≥10). Aligned reads were quantified according to UCSC (25) RefGene annotations for hg19 using LifeScope software.

Classification models and data visualization

RNA-seq expression data was analysed and visualized as log transformed (base 10) or standardized (z-score) Reads Per Kilobase of transcript per Million mapped reads (RPKM) values. Gene expression values from qPCR were analysed and visualized as –ΔΔCq values. Microarray gene expression data were analysed and visualized as Cy5/Cy3 intensity ratios. All gene expression values were standardized prior to being used in the support vector machine (SVM) models. In order to use the different gene expression measures interchangeably in a single model, a novel method was developed to transform the expression values of into new features. In brief, the expression values of the marker genes (e.g., IL36G, CCL27, NOS2 and C10orf99) as RPKM values, –ΔΔCq values or Cy5/Cy3 intensity ratios were transformed into polynomial equation coefficients followed by z-score standardization. This method is limited to be used with a small number of genes. The transformation method is explained in detail in Appendix S1[1].

SVM models were implemented using the e1071 package (26) in R. Gamma and cost parameters were determined by the *tune.svm()* function using values of 2^x, where x was an integer in the range [–15, 3] for gamma and [–5, 15] for cost parameters. Linear (1 feature) or sigmoid (more than 1 feature) kernels were used in the SVM models. Leave-one-out method was used for cross-validation (LOOCV) of the trained models. ROCR package (27) was used to plot receiver operating characteristic (ROC) curves and calculate the area under the curve (AUC).

Feature selection was conducted using the Recursive Feature Elimination (RFE) method from the caret package (28) in order to determine the best 4-gene sets for PsV classification. RFE was implemented using the predefined Random Forest function with 10-fold cross-validation. RFE was conducted separately on the GSE63741 and GSE121212 datasets and was limited to the genes present in both (n=1,419).

ggplot2 package (29) or R base graphics (30) were used for data visualization. Principal components were calculated using *prcomp()* function from the Stats package (30). Spearman's correlation coefficients were calculated to assess correlations between gene expression levels and patient characteristics for the PsV samples from both qPCR datasets and AD samples from the qPCR dataset 2.

Experimental design

Literature review for identification of proposed biomarkers. A review of published studies was conducted to identify potential biomarkers for PsV. The search was performed using the PubMed search engine (https://www.ncbi.nlm.nih.gov/pubmed) and the

[1]https://www.medicaljournals.se/acta/content/abstract/10.2340/00015555-3337

Table I. Datasets comprising the test set and corresponding number of samples

GEO[a]	Method[b]	PsV[c]	Ps-NL[d]	AD[e]	AD-NL[f]	LP[g]	CE[h]	C[i]	Ps-PP[j]	Ps-S[k]
GSE66511	RNA-seq	12	12					12		
GSE41745	RNA-seq	3	3							
GSE65832	RNA-seq			20	20					
GSE83645	RNA-seq	18	5							
GSE117405	RNA-seq	8						9	3	8
GSE121212	RNA-seq	28	27	27	27			38		
GSE63741	Microarray	30		30		30	30	30		

[a]NCBI GEO accession number. [b]Gene expression quantification method. [c]Lesional psoriasis vulgaris skin; [d]non-lesional skin of psoriasis patients; [e]lesional atopic dermatitis skin; [f]non-lesional skin of atopic dermatitis patients; [g]lesional lichen planus skin; [h]lesional contact dermatitis skin; [i]healthy control skin; [j]lesional plaque-type palmoplantar psoriasis skin; [k]lesional scalp psoriasis skin.

keyword "psoriasis biomarker". All review and research articles published in the years 2012 to 2017 were reviewed, including relevant research articles cited by the review articles. The full list of reviewed publications is shown in Table SI[1].

Validation of selected biomarker candidates. The selected gene expression-based biomarker candidates were validated in 2 steps. First, the differential expression of the marker genes was confirmed in RNA-seq dataset (GSE66511) of locally recruited patients and healthy controls followed by additional quantification by qPCR. Secondly, the disease-specificity of the selected markers was assessed using the GSE63741 and GSE121212 datasets and additionally recruited 20 patients with PsV and 18 patients with AD.

Classification model construction and testing. Five additional public datasets were included and classification models trained on qPCR data of locally recruited patients were validated on a total of 7 public datasets comprised of different patient groups. As an alternative data-driven biomarker selection approach, the GSE63741 and GSE121212 datasets were used to select 4-gene combinations that best differentiate PsV from other samples. Similarly, these multicomponent biomarkers were then validated on the full test set.

A graphical overview of the study design is presented in Fig. S1; Appendix S2[1]).

RESULTS

Biomarker candidates from the literature

From the review of published studies reporting biomarkers of PsV, it became evident that the expression levels of specific genes in the lesional skin hold the most promise for disease-specific biomarkers. In contrast, circulatory biomarkers seem to reflect more general inflammatory processes (14). As a result, 5 genes (*IL36G, NOS2, CCL27, C10orf99* and *IGFL1*) with altered expression levels in PsV lesions, which had previously been proposed as biomarkers of PsV by 3 independent studies, were selected to form the multicomponent biomarker (Table SII, Appendix S2[1]).

Confirmation of differential expression of the selected genes

The differential expression of the selected genes was confirmed in a previously published RNA-seq dataset (GSE66511) of 12 paired PsV skin samples and non-lesional skin samples from the same patients and 12 healthy controls. Despite the minor overlap of expression levels in PsV and Ps-NL or C samples in case of *IGFL1, NOS2* and *CCL27*, the differences in expression of the genes, as previously described, were apparent. Expression levels of the marker genes were further quantified by qPCR. The combined expression pattern of the 5 genes clearly distinguished PsV samples from Ps-NL and C skin samples (Fig. S2; Appendix S2[1]).

Disease-specificity of the selected genes and classification models

The disease-specificity of the selected biomarker genes and the classification models based on them were first assessed independently on the GSE63741, GSE121212, and combined qPCR datasets. It should be noted that the probes for one of the selected marker genes (*IGFL1*) were not included on the microarrays used to create the GSE63741 dataset and this was a limiting factor to the analysis.

The GSE63741 dataset is characterized by significantly higher *IL36G* expression levels in PsV samples compared with other inflammatory skin diseases and control samples (**Fig. 1**a), making it a near perfect biomarker for PsV, as reported by D'Erme and colleagues (31). None of the 3 other biomarker genes displayed comparable disease-specific qualities in this dataset, as *C10orf99* was upregulated in all the inflammatory skin diseases, *CCL27* was downregulated in both PsV and LP samples, and *NOS2* displayed only minor upregulation in all skin disease samples compared with controls. Leave-one-out cross-validation (LOOCV) of a SVM classification model based on *IL36G* expression yielded perfect accuracy (AUC=1) for this dataset. The *NOS2* and *CCL27* combination proposed by Quaranta and colleagues (5) resulted in a relatively poor classification model (AUC=0.84) in this case. A model based on all 4 genes performed with LOOCV accuracy comparable to *IL36G*-based classification (Fig. S3a–c; Appendix S2[1]).

Interestingly, in the combined qPCR datasets with 2 independent PsV groups, *IL36G* expression levels alone did not allow as accurate classification of PsV samples (AUC=0.95) due to the overlap of expression levels in the PsV and the AD groups (Fig. 1b). In the confines of this dataset, a SVM classification model based on *NOS2* and *CCL27* outperformed classification based on only *IL36G*. Classification accuracy further improved slightly with the 4-gene (excluding *IGFL1*) and 5-gene models. Similarly, the 4-gene and 5-gene models displayed better results compared with the *IL36G* or *NOS2* and *CCL27*-based models for the GSE121212 dataset due to overlapping expression levels in the different groups of samples (Fig. 1c and Fig. S3d–k; Appendix S2[1]).

The recurrent theme in the comparison of the 3 datasets was the consistent accuracy of sample classification by the 4-gene model (and the 5-gene model, where applicable), as the combined expression profile of these genes better survived the effect of inter-dataset variability (Fig. 1d–f). Therefore, the 4-gene model based on the expression levels of *L36G*, *CCL27*, *NOS2* and *C10orf99* with high LOOCV accuracies (95%) in all 3 independent datasets was selected as the preferred disease-specific classification model for PsV (Fig. S4; Appendix S2[1]).

In addition, the correlations between the marker genes and patient characteristics in the qPCR datasets were assessed, but no statistically significant ($p \leq 0.05$) correlations were observed between the expression levels of the marker genes and disease severity in either the PsV or the AD group (Table SIII; Appendix S2[1]).

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

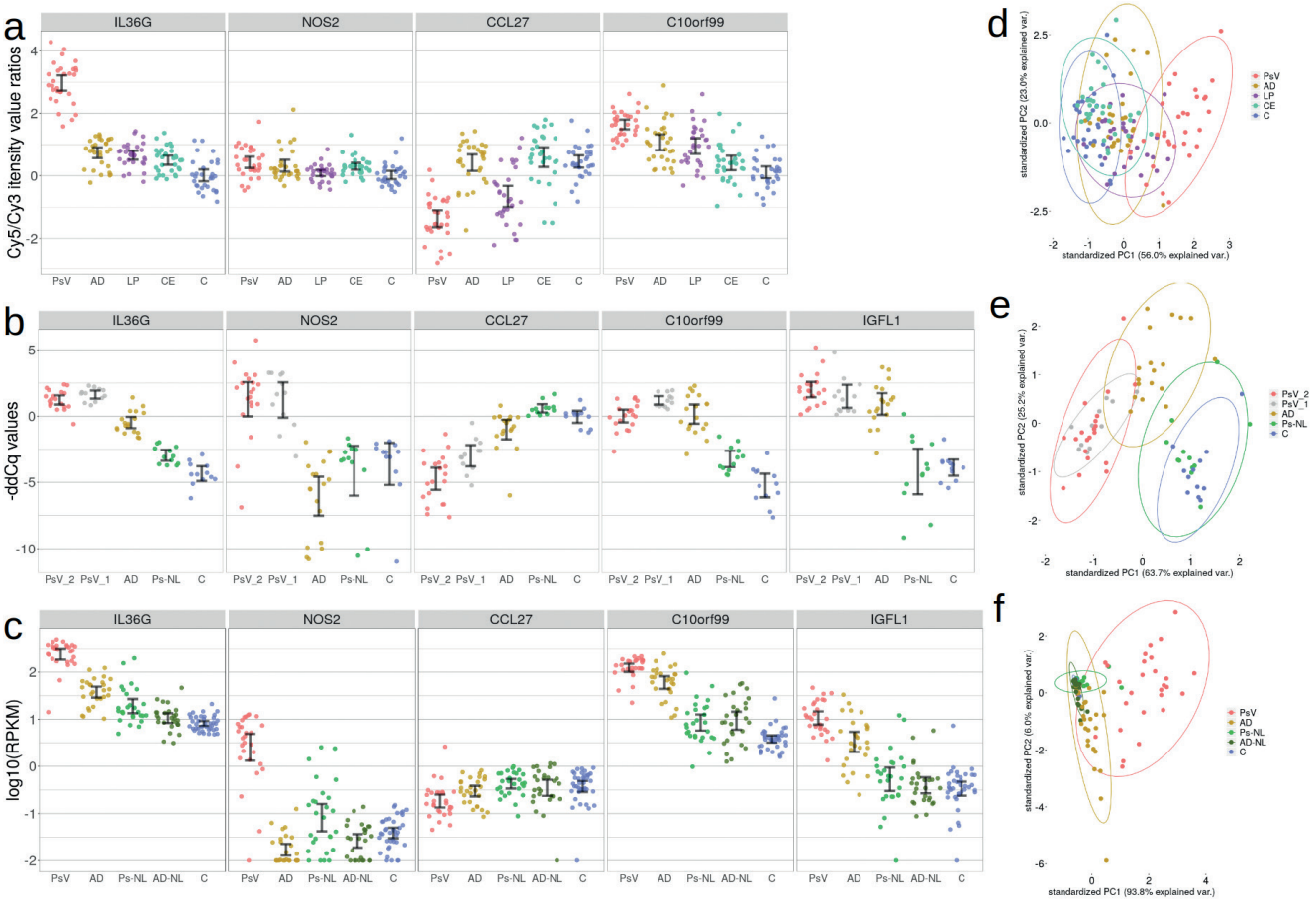


Fig. 1. Overview of marker gene expression levels in the GSE63741, GSE121212 and qPCR datasets. (a) Expression levels as Cy5/Cy3 intensity value ratios of the 4 out of 5 biomarker genes across 30 psoriasis (Ps), 30 atopic dermatitis (AD), 30 contact eczema (CE), 30 lichen planus (LP) and 30 healthy control (C) skin samples in the GSE63741 dataset. The expression levels of *IGFL1* were absent from this dataset. (b) Expression levels as –ddCq (ΔΔCq, quantitation cycle of the target gene adjusted for quantitation cycles of reference genes and calibration sample) values of the 5 marker genes for 32 psoriasis vulgaris (PsV_1 and PsV_2), 18 AD, 12 non-lesional psoriasis (Ps-NL) and 12 C skin samples in the qPCR datasets. (c) Expression levels as log10 of RPKM (Reads Per Kilobase of transcript per Million mapped reads) values of the 5 marker genes for 28 PsV, 27 Ps-NL, 27 AD and 38 C skin samples in the GSE121212 dataset. The first 2 principal components of the 4 (or 5) gene expression values in the (d) GSE63741, (e) qPCR, and (f) GSE121212 datasets. Error bars represent 95% confidence intervals.

Test set assembly and transformation of gene expression values

A test set comprised of 7 public datasets (Table I) totalling 504 samples was assembled to fully test the performance of the 4-gene classification model. Gene expression values in these datasets were transformed into new features to correct for the differences in values obtained by different gene expression quantification methods (RNA-seq, microarray, and qPCR). The transformation of expression values of the 4 marker genes (*IL36G*, *CCL27*, *NOS2* and *C10orf99*) into polynomial equation coefficients followed by z-score standardization rescaled the RPKM values, Cy5/Cy3 intensity ratios, and –ΔΔCq values, while retaining the overall expression profiles (**Fig. 2** and Figs S5–8; Appendix S2[1]). This enabled the interchangeable use of different gene expression values in a single model. Transformation into polynomial equation coefficients prior to z-score standardization was essential to rescaling the data and maintaining the expression pattern in the form of new features, and outperformed simple z-score standardization without transformation (Fig. S9; Appendix S2[1]).

Model testing

A new 4-gene model was trained on the transformed qPCR data with a LOOCV accuracy of 95.9%, comparable to the model trained on untransformed data (Fig. S10; Appendix S2[1]). Three PsV samples produced false-negative predictions during LOOCV. In these samples, the expression levels of *IL36G* and *NOS2* were considerably lower than the majority of the PsV samples (Fig. S11; Appendix S2[1]).

The model was then tested on an assembled set of 493 samples, labelled PsV or Ot (other). The accuracy of model predictions on the validation set was 96.4%, with 6 false-negative and 9 false-positive calls, with PsV as the positive outcome (**Fig. 3**). Falsely predicted samples in both the training and test set differentiate from their respective groups based on their gene expression profiles

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

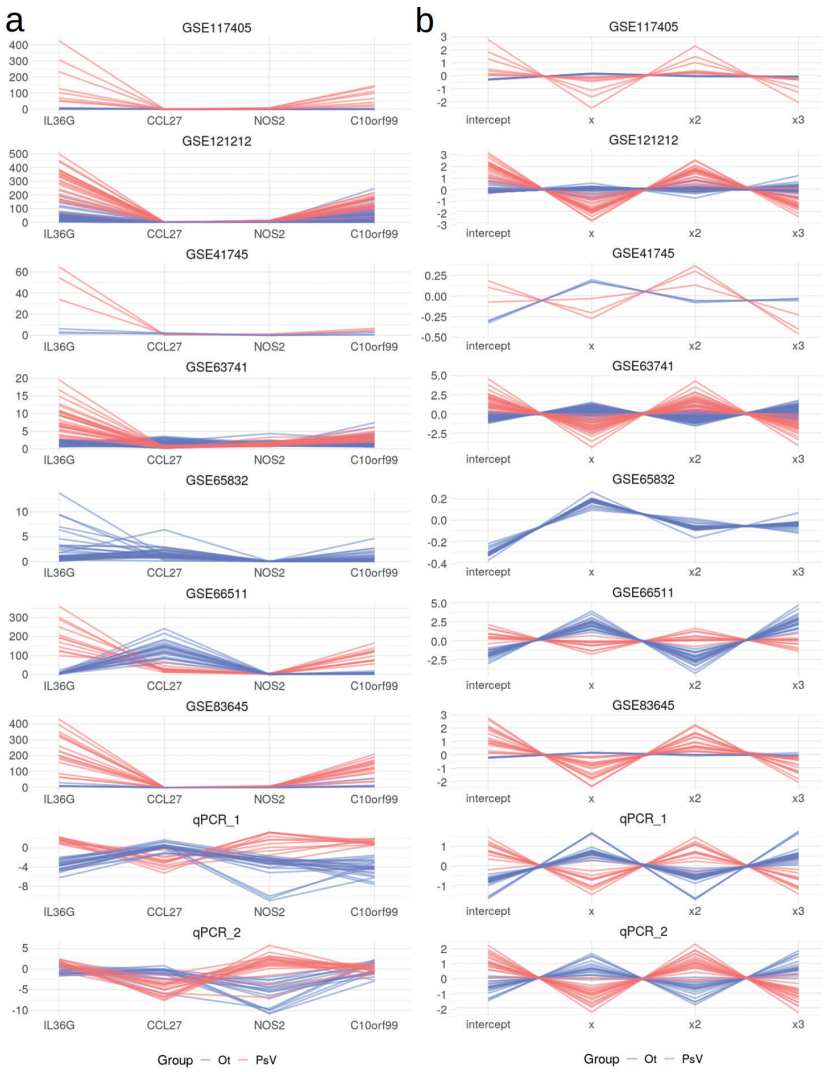


Fig. 2. Expression values of initial marker genes transformed to new features. (a) Untransformed expression values for the 4 marker genes as Reads Per Kilobase of transcript per Million mapped reads (RPKM) values for RNA-seq datasets (GSE117405, GSE121212, GSE41745, GSE65832, GSE66511, and GSE83645), Cy5/Cy3 intensity value ratios for the GSE63741 microarray dataset or −ΔΔCq values (quantitation cycle of the target gene adjusted for quantitation cycles of reference genes and calibration sample) for qPCR datasets. Samples are coloured on the graphs as lesional psoriasis vulgaris skin samples (PsV: *red*) or other samples (Ot: *blue*). The latter consists of non-lesional skin samples of psoriasis vulgaris patients, atopic dermatitis, non-lesional atopic dermatitis, contact eczema, lichen planus and healthy control skin samples. (b) Expression values for the 4 marker genes transformed into new features intercept, x, x2 and x3 as in $y=\beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \varepsilon_1$. The polynomial line equation was obtained by coding the 4 marker genes as integers and using a linear model to fit a degree 3 polynomial line to connect expression values along the x-axis. The new features were further standardized as z-scores, where standard deviation (SD) and mean were calculated separately for each gene and expression quantification method group (RNA-seq, microarray or qPCR) successfully rescaling the gene expression values obtained by different quantification methods.

(Fig. S12; Appendix S2[1]) and are listed in Table SIV; Appendix S2[1]. Classifiers based solely on *IL36G* expression levels or the combination of *NOS2* and *CCL27* genes, resulted in respective prediction accuracies of 94.7% and 74.7% on the test set (Fig. S13; Appendix S2[1]).

The SVM model was also used to classify 8 scalp psoriasis (Ps-S) and 3 plaque-type palmoplantar psoriasis (Ps-PP) samples from the GSE117405 dataset. Seven Ps-S samples were classified as PsV and all of the Ps-PP were classified as Ot. Ps-PP samples displayed *NOS2* expression levels comparable to the expression levels of healthy control samples from the same dataset, leading the model not to recognize them as PsV (Fig. S14; Appendix S2[1]).

Data-driven alternative to biomarker candidate selection

In addition to the set of marker genes selected based on the results of previously published studies, a data-driven biomarker selection was conducted using the GSE63741 and GSE121212 datasets. The RFE method was used to select the best 4-gene combinations differentiating PsV samples in both datasets. Inherent to the GSE63741 dataset, *IL36G* proved to be the best PsV-specific biomarker and including the 3 next best genes, *CRABP2*, *S100A7A*, and *IL36RN*, marginally increased resulting prediction errors (Fig. S15a; Appendix S2[1]). In case of the GSE121212 dataset, prediction errors decreased with each added marker gene (Fig. S15b; Appendix S2[1]). The top 4 predictors of PsV based on the GSE121212 dataset were *SPRR2A, PRELP, ARG1* and *KYNU*. The expression levels of these 8 genes were mostly consistent between the 2 datasets, with the exception of *PRELP*, which was downregulated in PsV compared with AD samples in the GSE121212, but not in the GSE63741 dataset (**Fig. 4** a–b).

Validated on their respective source datasets, the LOOCV accuracies of the SVM classification models based on these multicomponent biomarkers were 98.7% and 95.9% for the GSE63741 and GSE121212 gene set, respectively (Fig. 4c–d). The accuracy of the classification model trained on the GSE63741 set and tested on GSE121212 was 95.9%. The accuracy of the model trained and tested vice versa was 90.0% (Fig. 4e–f). When tested on the full transformed test set, the performances of both SVM models were inferior to the initial multicomponent biomarker (*IL36G, CCL27, NOS2* and *C10orf99*) with accuracies of 88.8% and 88.6% for the GSE63741-derived gene set and the GSE121212-derived gene set, respectively (Fig. S16; Appendix S2[1]). The previously described method of data transformation was applied to construct the test set for these models (Figs S17–20; Appendix S2[1]).

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

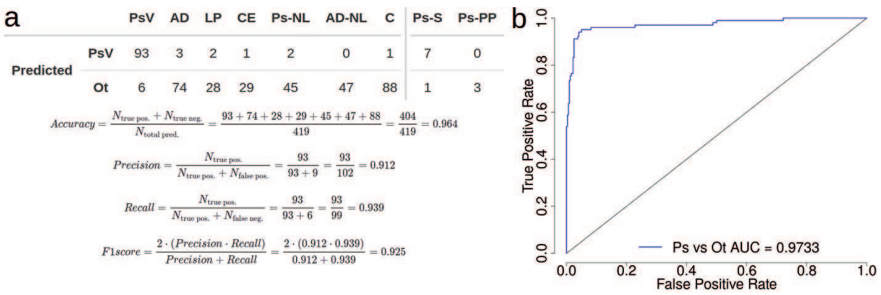
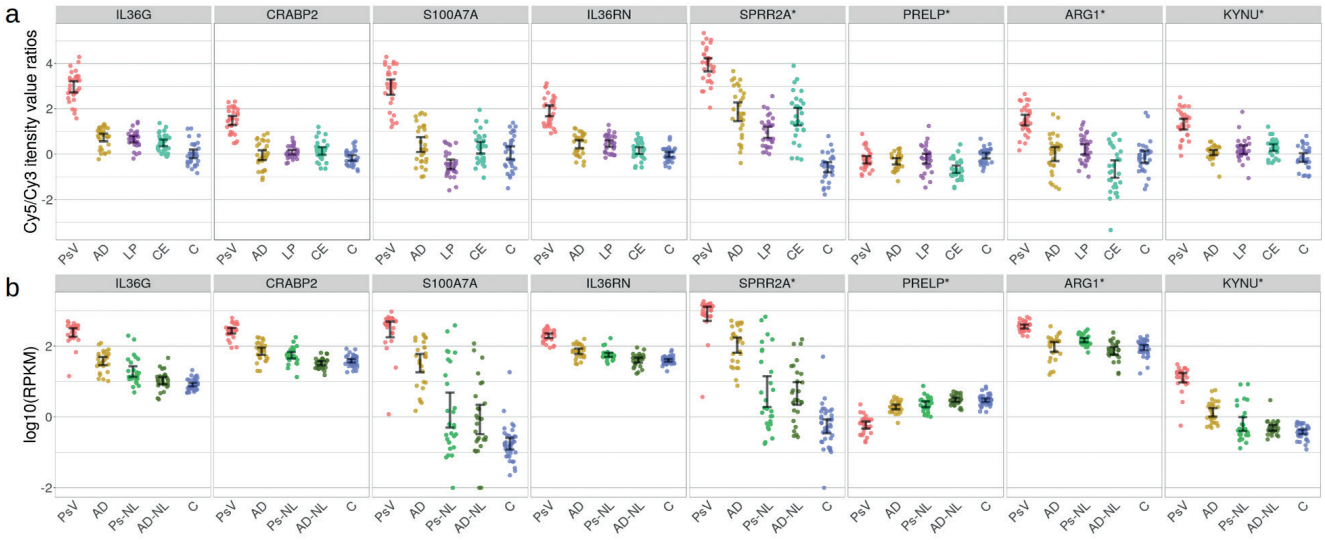


Fig. 3. Support vector machine (SVM) classification model performance on transformed test set. (a) Predictions of the SVM classification model trained on transformed qPCR datasets and tested on 6 transformed RNA-seq and 1 transformed microarray datasets of 99 psoriasis vulgaris (PsV), 47 non-lesional psoriasis (Ps-NL), 77 atopic dermatitis (AD), 47 non-lesional atopic dermatitis (AD-NL), 30 lichen planus (LP), 30 contact eczema (CE), and 88 C skin samples. The predictions on scalp psoriasis (Ps-S) and palmoplantar psoriasis (Ps-PP) samples were not included in the calculations of different model performance metrics. (b) Receiver operating characteristic (ROC) plot and corresponding area under the curve (AUC) value based on the prediction scores obtained from testing the model trained on qPCR data on the assembled test set.

DISCUSSION

Several gene expression-based biomarkers for PsV have been proposed, but none of them have been completely discriminatory for PsV or confirmed in multiple patient cohorts. In the present study, it was hypothesized that using a classification model based on several potential biomarkers proposed by separate studies with independent patient cohorts would ensure improved results in the biomarker-based diagnosis of a heterogeneous disease, such as PsV. As such, the aim of this study was to present a proof of concept classification model capable of more accurate classification of PsV samples across different patient groups than any of the marker genes individually. A SVM classification model was trained based on the expression levels of 4 genes (*IL36G, CCL27, NOS2* and *C10orf99*) in the lesional skin tissue using 2 independent sets of samples from locally recruited patients. The model was validated on 7 different public datasets, resulting in a 96.4% accuracy of distinguishing PsV from other inflammatory skin diseases. This multicomponent biomarker combination outperformed classification based on single marker genes and other gene sets, which were more affected by the variability of gene expression profiles between different datasets and patient groups.

This study was limited by the availability of (gene expression) data from large patient cohorts and the inter-dataset variability inherent to the available datasets. Conventional application of machine learning methods implies large datasets, the absence of which can lead



c

Pred.	PsV	AD	LP	CE	C
PsV	28	0	0	0	0
Ot	2	30	30	30	30

d

Pred.	PsV	AD	Ps-NL	AD-NL	C
PsV	24	1	1	0	0
Ot	4	26	26	27	38

e

Pred.	PsV	AD	Ps-NL	AD-NL	C
PsV	23	0	1	0	0
Ot	5	27	26	27	38

f

Pred.	PsV	AD	LP	CE	C
PsV	24	4	2	3	0
Ot	6	26	28	27	30

Fig. 4. Overview of expression levels of the genes selected by the data-driven approach in the GSE63741 and GSE121212 datasets. (a) Expression levels as Cy5/Cy3 intensity value ratios of the 8 biomarker genes across 30 psoriasis (Ps), 30 atopic dermatitis (AD), 30 contact eczema (CE), 30 lichen planus (LP) and 30 healthy control (C) skin samples in the GSE63741 dataset. (b) Expression levels as log10 of RPKM (Reads Per Kilobase of transcript per Million mapped reads) values of the 8 marker genes for 28 PsV, 27 Ps-NL, 27 AD and 38 C skin samples in the GSE121212 dataset. The genes that are not marked by an asterisk were selected based on Recursive Feature Elimination (RFE) on the GSE63741 dataset and the genes that are marked by an asterisk were selected as the result of (RFE) on the GSE121212 dataset. Error bars represent 95% confidence intervals. (c) Results of leave-one-out cross-validation (LOOCV) of the classification model based on the GSE63741-derived gene set and (d) results of LOOCV of the GSE121212-derived gene set model. (e) Results of testing the classification model based on the GSE63741-derived gene set on the GSE121212 dataset and (f) results of testing of the GSE121212-derived gene set model on the GSE63741 dataset.

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

to limited accuracy and high possibility of overfitted models trained on small datasets (32–34). Public gene expression datasets of inflammatory skin diseases other than psoriasis are scarce and recruiting a large number of patients with different inflammatory skin diseases for a conventional machine learning application would be an extensively laborious task. Therefore, in this study, support vector machine (SVM) classification models were used in an unconventional setting with emphasis on training and testing the model on data from different groups of patients.

Defining the best possible combination of biomarkers for PsV diagnostics is heavily dependent on the available datasets and the possibility to integrate them. The availability of 2 larger datasets (GSE63741 and GSE121212) and the novel method developed as a part of this study to normalize different gene expression measures allowed the possibility to pursue this direction. RFE was used to find combinations of genes best differentiating PsV from other samples, similarly to the study by Guo et al. (35), but with the exception of including other inflammatory skin disease samples as well. Although this analysis was limited to the genes present in the GSE63741 microarray dataset, it did provide data-driven alternatives to the hypothesis-based selection of the genes making up the multicomponent biomarker.

Perhaps the most important observation in this study was that the expression profiles of individual marker genes were notably variable across different datasets. In the GSE63741 dataset, the expression levels of *IL36G* were uniformly different in PsV compared with AD skin samples, but displayed an overlap of expression values in the qPCR and GSE121212 datasets. Interestingly, in case of the 3 outlying PsV samples in the qPCR dataset in terms of lower *IL36G* expression levels, the lesions of these patients exhibited hallmark signs of PsV and were in the process of acute exacerbation during the time of sampling. Similarly, the expression levels of *NOS2* and *CCL27*, a gene pair that has previously been used to differentiate PsV from eczema samples (5), were surprisingly variable across the datasets included in this study and did not result in highly accurate predictions when used in a SVM classification model. For example, in the GSE63741 dataset, *NOS2* displayed comparable expression levels in PsV, AD, and CE samples and *CCL27* was downregulated in both PsV and LP samples. Conversely, in the confines of the qPCR datasets, expression levels of *NOS2* and *CCL27* in the PsV samples compared with AD samples were in line with the result reported by Quaranta and colleagues (5). From amongst the genes selected based on RFE, *PRLEP* was downregulated in PsV samples in the GSE121212 and GSE66511 datasets and improved the classification accuracy, while in the GSE63741 and GSE83645 datasets this was not the case.

Many factors could potentially lead to considerable differences between individual datasets (groups of patients), e.g. differences in patient characteristics, disease severity, cellular composition of the skin biopsy (36), genetic background (37, 38), non-representative sample sizes, or the effects introduced by the use of different expression quantification methods. Nevertheless, it was apparent throughout the 3 independent datasets used for LOOCV that the 4-gene (*IL36G, CCL27, NOS2* and *C10orf99*) multicomponent biomarker retained its accuracy, while individual genes displayed inconsistent expression levels. The combined expression profile of the genes making up the multicomponent biomarker was less affected by differences between datasets (groups of patients) compared with individual genes, demonstrating that the former to be a more consistent biomarker.

The performances of all constructed classification models were tested on the assembled set of 493 samples, which consisted of several inflammatory skin disease samples in addition to PsV and control samples (Table I). As this was an imbalanced dataset, with 99 of the 493 being PsV samples, the combined measure of precision and recall, the F1 score, is best suited to compare the different models (39). The 4-gene (*IL36G, CCL27, NOS2* and *C10orf99*) model yielded F1 score of 0.925 (96.4% accuracy) and the next-best model based only on *IL36G* resulted in a F1 score of 0.880. Classification based on *IL36G* yielded a recall value of 0.818 compared with the 0.939 of the 4-gene model, resulting in a less sensitive, yet slightly more precise model. Due to the inter-dataset variation of *NOS2* and *CCL27* expression levels, classification based on these genes was poorer, with a F1 score of 0.597, falling short of the level of performance reported by Quaranta and colleagues. The performance of this classifier was heavily affected by the expression levels of *NOS2* in the GSE63741 dataset. The RFE on GSE63741 and GSE121212 datasets yielded models with respective F1 scores of 0.732 and 0.780. Therefore, the data-driven selection of marker genes for the multicomponent model did not result in a better combination of genes than the initial hypothesis-based selection.

In conclusion, combining multiple gene expression markers as a multicomponent biomarker allowed the distinction of PsV from other inflammatory skin diseases more accurately than any of the markers individually. This approach could potentially be further expanded to differentiate between disease subtypes (40), improve distinction in cases of overlap of different inflammatory skin disorders, and to more accurately monitor the efficacy of treatments. However, the need for validation in larger groups of patients remains. Coupled with a cost-effective gene expression quantification technique, this approach has the potential to fill the void of clinically applicable biomarkers for PsV and other inflammatory skin diseases.

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

ACKNOWLEDGEMENTS

This work was supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 668989, by the personal research Grant PUT1465 (from the Estonian Research Council), by institutional research funding IUT20-46 (from the Estonian Ministry of Education and Research), and European Regional Development Fund.

Datasets related to this article (Table I) can be found in the NCBI GEO repository (https://www.ncbi.nlm.nih.gov/geo/).

The authors have no conflicts of interest to declare.

REFERENCES

1. Christophers E. Psoriasis – epidemiology and clinical spectrum. Clin Exp Dermatol 2001; 26: 314–320.
2. Griffiths CE, Barker JN. Pathogenesis and clinical features of psoriasis. Lancet 2007; 370: 263–271.
3. Bordignon M, Belloni-Fortina A, Pigozzi B, Saponeri A, Alaibac M. The role of immunohistochemical analysis in the diagnosis of parapsoriasis. Acta Histochem 2011; 113: 92–95.
4. Guttman-Yassky E, Krueger JG. Atopic dermatitis and psoriasis: two different immune diseases or one spectrum? Curr Opin Immunol 2017; 48: 68–73.
5. 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.
6. Smith CH, Barker JNWN. Psoriasis and its management. BMJ 2006; 333: 380–384.
7. Nowicki R, Trzeciak M, Wilkowska A, Sokołowska-Wojdyło M, Ługowska-Umer H, Barańska-Rybak W, et al. Atopic dermatitis: current treatment guidelines. Statement of the experts of the Dermatological Section, Polish Society of Allergology, and the Allergology Section, Polish Society of Dermatology. Adv Dermatol Allergol 2015; 32: 239–249.
8. Saary J, Qureshi R, Palda V, DeKoven J, Pratt M, Skotnicki-Grant S, et al. A systematic review of contact dermatitis treatment and prevention. J Am Acad Dermatol 2005; 53: 845.
9. Atzmony L, Reiter O, Hodak E, Gdalevich M, Mimouni D. Treatments for cutaneous lichen planus: a systematic review and meta-analysis. Am J Clin Dermatol 2016; 17: 11–22.
10. Golińska J, Sar-Pomian M, Rudnicka L. Dermoscopic features of psoriasis of the skin, scalp and nails – a systematic review. J Eur Acad Dermatol Venereol 2019; 33: 648-660.
11. Jha AK, Lallas A, Sonthalia S, Jhakar D, Udayan UK, Chaudhary RKP. Differentiation of pityriasis rubra pilaris from plaque psoriasis by dermoscopy. Dermatol Pract Concept 2018; 8: 299–302.
12. Lallas A, Kyrgidis A, Tzellos TG, Apalla Z, Karakyriou E, Karatolias A, et al. Accuracy of dermoscopic criteria for the diagnosis of psoriasis, dermatitis, lichen planus and pityriasis rosea. Br J Dermatol 2012; 166: 1198–1205.
13. Cook LC, Hanna C, Foulke GT, Seiverling EV. Dermoscopy in the diagnosis of inflammatory dermatoses: systematic review findings reported for psoriasis, lupus, and lichen planus. J Clin Aesthetic Dermatol 2018; 11: 41–42.
14. Jiang S, Hinchliffe TE, Wu T. Biomarkers of an autoimmune skin disease – psoriasis. Genomics Proteomics Bioinformatics 2015; 13: 224–233.
15. Coffman E, Richmond-Bryant J. Multiple biomarker models for improved risk estimation of specific cardiovascular diseases related to metabolic syndrome: a cross-sectional study. Popul Health Metr 2015; 13: 7.
16. Shariat SF, Karakiewicz PI, Ashfaq R, Lerner SP, Palapattu GS, Cote RJ, et al. Multiple biomarkers improve prediction of bladder cancer recurrence and mortality in patients undergoing cystectomy. Cancer 2008; 112: 315–325.
17. Cretu D, Gao L, Liang K, Soosaipillai A, Diamandis EP, Chandran V. Differentiating psoriatic arthritis from psoriasis without psoriatic arthritis using novel serum biomarkers. Arthritis Care Res 2018; 70: 454–461.
18. Chen D, Jain S, Su Y-H, Song W. Building classification models with combined biomarker tests: application to early detection of liver cancer. J Stat Sci Appl 2017; 5: 91–103.
19. Lazzarini N, Runhaar J, Bay-Jensen AC, Thudium CS, Bierma-Zeinstra SMA, Henrotin Y, et al. A machine learning approach for the identification of new biomarkers for knee osteoarthritis development in overweight and obese women. Osteoarthritis Cartilage 2017; 25: 2014–2021.
20. Mazzara S, Rossi RL, Grifantini R, Donizetti S, Abrignani S, Bombaci M. CombiROC: an interactive web tool for selecting accurate marker combinations of omics data. Sci Rep 2017; 7: 45477.
21. Keermann M, Kõks S, Reimann E, Prans E, Abram K, Kingo K. Transcriptional landscape of psoriasis identifies the involvement of IL36 and IL36RN. BMC Genomics 2015; 16: 322.
22. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 2009; 55: 611–622.
23. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinforma Oxf Engl 2013; 29: 15–21.
24. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res 2012; 22: 1760–1774.
25. Meyer LR, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, Wong M, et al. The UCSC Genome Browser database: extensions and updates 2013. Nucleic Acids Res 2013; 41: D64-69.
26. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.6-8.
27. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinforma Oxf Engl 2005; 21: 3940–3941.
28. Kuhn M. caret: Classification and regression training. R package version 6.0-84.
29. Wickham H. ggplot2: Elegant graphics for data analysis. R package version 3.2.0.
30. R Core Team. R: A language and environment for statistical computing. R version 3.6.0.
31. D'Erme AM, Wilsmann-Theis D, Wagenpfeil J, Hölzel M, Ferring-Schmitt S, Sternberg S, et al. IL-36γ (IL-1F9) is a biomarker for psoriasis skin lesions. J Invest Dermatol 2015; 135: 1025–1032.
32. Figueroa RL, Zeng-Treitler Q, Kandula S, Ngo LH. Predicting sample size required for classification performance. BMC Med Inform Decis Mak 2012; 12: 8.
33. Popovici V, Chen W, Gallas BG, Hatzis C, Shi W, Samuelson FW, et al. Effect of training-sample size and classification difficulty on the accuracy of genomic predictors. Breast Cancer Res 2010; 12: R5.
34. Bzdok D, Krzywinski M, Altman N. Machine learning: a primer. Nat Methods 2017; 14: 1119–1120.
35. Guo P, Luo Y, Mai G, Zhang M, Wang G, Zhao M, et al. Gene expression profile based classification models of psoriasis. Genomics 2014; 103: 48–55.
36. Kim BY, Choi JW, Kim BR, Youn SW. Histopathological findings are associated with the clinical types of psoriasis but not with the corresponding lesional psoriasis severity index. Ann Dermatol 2015; 27: 26–31.
37. Marquand AF, Rezek I, Buitelaar J, Beckmann CF. Understanding heterogeneity in clinical cohorts using normative models: beyond case-control studies. Biol Psychiatry 2016; 80: 552–561.
38. Magosi LE, Goel A, Hopewell JC, Farrall M, Consortium on behalf of the Cardi. Identifying systematic heterogeneity patterns in genetic association meta-analysis studies. PLOS Genet 2017; 13: e1006755.
39. Tharwat A. Classification assessment methods. Appl Comput Inform 2018; published online.
40. Ahn R, Yan D, Chang H-W, Lee K, Bhattarai S, Huang Z-M, et al. RNA-seq and flow-cytometry of conventional, scalp, and palmoplantar psoriasis reveal shared and distinct molecular pathways. Sci Rep 2018; 8: 11368.