

<https://helda.helsinki.fi>

The integration of large-scale public data and network analysis uncovers molecular characteristics of psoriasis

Federico, Antonio

2022-11-28

Federico , A , Pavel , A , Möbus , L , McKean , D , Del Giudice , G , Fortino , V , Niehues , H , Rastrick , J , Eyerich , K , Eyerich , S , van den Bogaard , E , Smith , C , Weidinger , S , de Rinaldis , E & Greco , D 2022 , ' The integration of large-scale public data and network analysis uncovers molecular characteristics of psoriasis ' , Human Genomics , vol. 16 , no. 1 , 62 . <https://doi.org/10.1186/s40246-022-00431-x>

<http://hdl.handle.net/10138/352022>

<https://doi.org/10.1186/s40246-022-00431-x>

cc_by

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

RESEARCH

Open Access



The integration of large-scale public data and network analysis uncovers molecular characteristics of psoriasis

Antonio Federico^{1,2,3,4}, Alisa Pavel^{1,2,3}, Lena Möbus^{1,2,3}, David McKean⁵, Giusy del Giudice^{1,2,3}, Vittorio Fortino⁶, Hanna Niehues⁷, Joe Rastrick⁸, Kilian Eyerich^{9,10}, Stefanie Eyerich¹¹, Ellen van den Bogaard⁷, Catherine Smith¹², Stephan Weidinger¹³, Emanuele de Rinaldis⁵ and Dario Greco^{1,2,3,14*}

Abstract

In recent years, a growing interest in the characterization of the molecular basis of psoriasis has been observed. However, despite the availability of a large amount of molecular data, many pathogenic mechanisms of psoriasis are still poorly understood. In this study, we performed an integrated analysis of 23 public transcriptomic datasets encompassing both lesional and uninvolved skin samples from psoriasis patients. We defined comprehensive gene co-expression network models of psoriatic lesions and uninvolved skin. Moreover, we curated and exploited a wide range of functional information from multiple public sources in order to systematically annotate the inferred networks. The integrated analysis of transcriptomics data and co-expression networks highlighted genes that are frequently dysregulated and show aberrant patterns of connectivity in the psoriatic lesion compared with the unaffected skin. Our approach allowed us to also identify plausible, previously unknown, actors in the expression of the psoriasis phenotype. Finally, we characterized communities of co-expressed genes associated with relevant molecular functions and expression signatures of specific immune cell types associated with the psoriasis lesion. Overall, integrating experimental driven results with curated functional information from public repositories represents an efficient approach to empower knowledge generation about psoriasis and may be applicable to other complex diseases.

Keywords: Psoriasis, Transcriptomics, Network analysis, Biomarkers, Public data, Druggability

Background

Psoriasis is a chronic inflammatory disorder of the skin, characterized by abnormal keratinocyte differentiation and hyper-proliferation of the epidermis, along with infiltration of inflammatory cells [1]. Although genetic [2] and environmental factors are known to contribute to the etiology of this polygenic disease, many of the intricate mechanisms of molecular alteration underlying the disease remain largely uncovered [3]. Multiple

transcriptome studies have pinpointed key pathways altered in lesional psoriasis skin [4–8]. However, integrated analysis of multiple homogenized datasets is, to date, still limited to a few examples [9, 10]. Although, for instance, Piruzian and colleagues [9] report the results of an integrated meta-analysis of both protein and gene expression datasets (and, therefore, integrating different data types), they still summarize the results of single datasets. Moreover, the analytical strategies employed in such studies have an impact on the ability to disentangle more complex patterns of molecular deregulation. In fact, while univariate differential expression analysis shed light on hundreds (sometimes thousands) of dysregulated genes in the lesional skin,

*Correspondence: dario.greco@tuni.fi

¹ Faculty of Medicine and Health Technology, Tampere University, Kauppi Campus, Arvo Ylpön Katu 34, 33520 Tampere, Finland
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

it is not straightforward to infer regulatory loops of molecular alterations underlying the phenotype.

This gap of knowledge could be filled by exploiting the large amount of biological data accumulated in recent years. In fact, vast amounts of data have been collected in public repositories and made freely available to the scientific community.

However, integrating such a wealth of data sources is still challenging due to the heterogeneity of data formats and the need for extensive manual curation [11].

A rigorous integration and exploitation of public data can provide a double benefit. On the one hand, already available data can inform the design of novel experimental strategies in order to achieve new knowledge. On the other hand, publicly available data may provide a shortcut to characterize and interpret *de novo* findings derived from targeted experiments. In the context of psoriasis, several repositories, such as Pharos (<https://pharos.nih.gov/>) [12], Target Validation [<https://www.targetvalidation.org>], Human Protein Atlas (<https://www.proteinatlas.org>) [13] and Clinical Trials (<https://www.clinicaltrials.gov>), report fundamental information about the state of the art of research in this topic, starting from the druggability/tractability of suitable drug targets to clinical trials and large-scale genetic association studies. Such data have never been integrated in order to derive new knowledge about the mechanistic events underlying the psoriatic phenotype.

Graph theory provides effective models to uncover the relevant gene–gene expression relationships both in physiological and in pathological conditions [14, 15]. In fact, gene co-expression network analysis is currently employed to understand the relationship between pairs of genes, and ultimately, gene networks or modules representing a marker of impaired biological functions in a disease [16].

In this study, we have integrated gene expression analysis and co-expression network analysis approaches on 23 manually curated transcriptomics datasets [17] in order to (1) validate and prioritize genes that are already known to be associated with psoriasis, (2) uncover novel genes never associated with psoriasis before, and (3) create a gene-centric compendium of psoriasis-related information curated from multiple data repositories. The whole

analytical pipeline implemented for this study is shown in Fig. 1.

Methods

Data collection and preprocessing

All the raw transcriptomics data collected for this study are publicly available in the Gene Expression Omnibus (GEO) repository. The data consist of 23 microarray-derived gene expression datasets of both lesional (574 samples) and non-lesional skin (540 samples) of psoriasis patients. GEO IDs of the collected datasets are reported in Additional file 1: Table S1.

As described in Federico et al. [17], the preprocessing of collected transcriptomics datasets was performed through the use of the eUTOPIA software [18], which implements all of the following steps and the relative functions. *Quality check*: for Agilent datasets, each sample was evaluated by visual inspection of the array pseudo-images, quality check reports and density plots of probe intensities. Therefore, outlier samples were removed from the analysis. For the Affymetrix datasets, outlier samples were detected by computing the Normalized Unscaled Standard Error (NUSE) [19] and the Relative Log Expression (RLE) [19] from the affyPLM v1.64.0R package, and the RNA degradation curves (RNADeg) from the affy v1.64.0R package [20]. The distributions of the values of these three metrics were investigated by means of boxplots and the sample outlieriness was evaluated for each measure based on the data distribution. Eventually, a concordance outlieriness score was computed across the three metrics. In particular, a sample was removed from the analysis if considered an outlier in at least two out of three metrics, one of them being the RNA degradation curve. Affymetrix-based studies were normalized by using the justRMA from the R affy v1.66.0 package [20]. Agilent-based studies were quantile normalized with the `normalizeQuantiles` function from the `limma` v3.44.3 package [21]. In order to investigate the effect of unknown batches that might mask biological variability, surrogate variable analysis (SVA) was performed with the `eUtopia` software, which implements the `sva` R package [22]. The analysis was performed by using the status of the skin (lesional and non-lesional) as variable of interest. The other biological variables (if present and if not confounded with the variable of interest)

(See figure on next page.)

Fig. 1 Overview of the analytical pipeline conceived and developed in this study. **A** Integration of transcriptomics public datasets through a gene expression integrated analysis. Differentially expressed genes reported in all of the included DNA microarray platforms were used in order to build gene co-expression network models of lesional and non-lesional skin from psoriasis patients. **B** Network analyses performed on the inferred networks. In particular, the analyses carried out in this study include a differential centrality analysis between the lesional and non-lesional skin network models, the identification of bridge genes in the lesional network, the functional annotation of co-expression modules of the lesional network, the enrichment of immune cell-specific genes and the evaluation of the druggability of the lesional skin network. Panel C shows the psoriasis disease map inferred in the present study

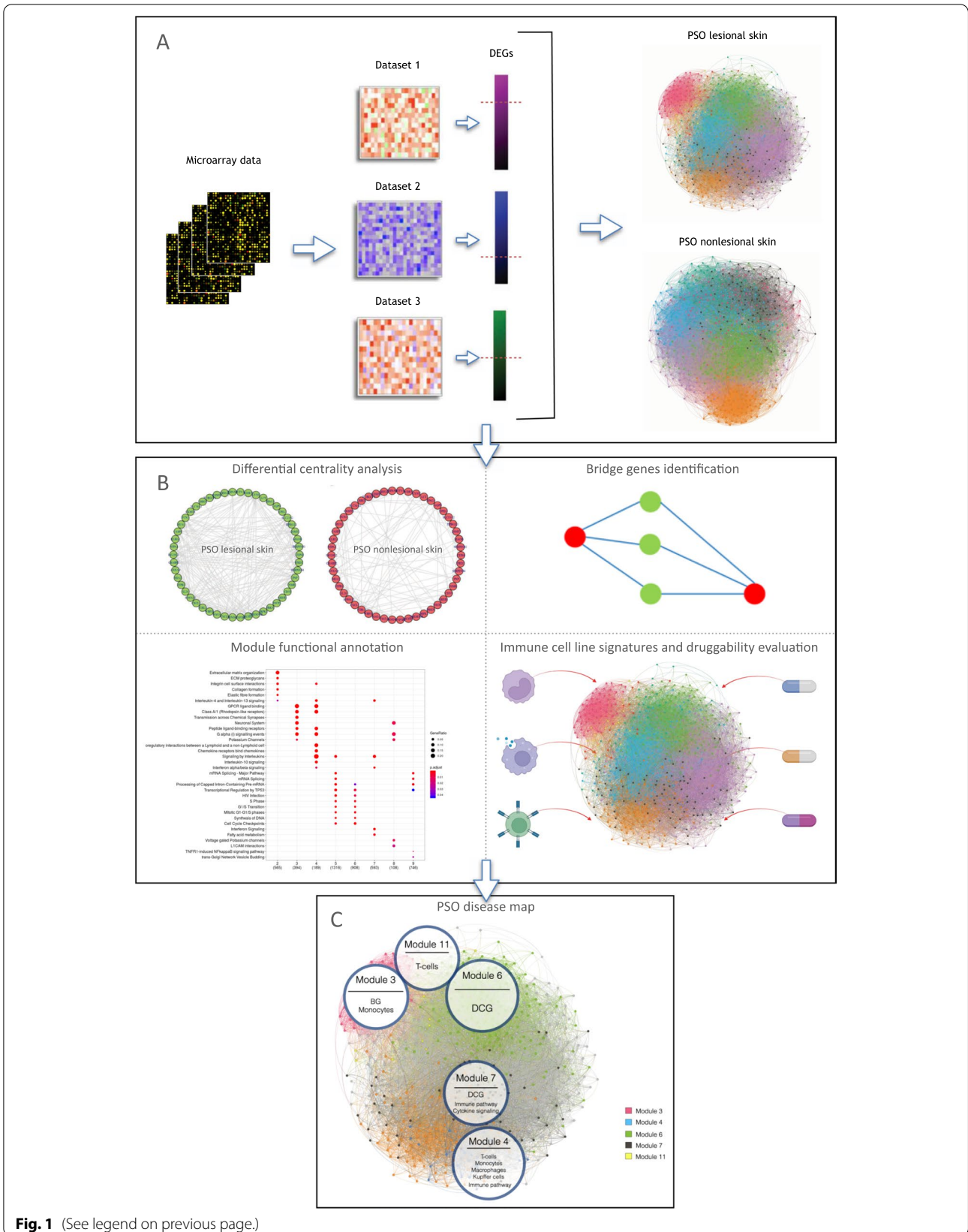


Fig. 1 (See legend on previous page.)

were used as covariates. Custom annotation files (CDF files) were downloaded from Brainarray (http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/CDF_download.asp) for Affymetrix-based microarrays. The latest version of Agilent probe annotation was retrieved from the Agilent Web site (<https://earray.chem.agilent.com/earray/>). The probesets were mapped to the Ensembl gene IDs, and the expression matrix was aggregated by computing the median of the expression of the Agilent probes mapping to the same Ensembl transcript ID. Only genes that are common to all the platforms were included in the analysis. Differentially expressed genes for each dataset were identified through the use of the limma package by comparing the lesional skin samples with the non-lesional skin. Resulting p values were corrected through the Benjamini–Hochberg method, and genes having an adjusted p value ≤ 0.05 were considered differentially expressed. Moreover, we assessed the consistency of deregulation of differentially expressed genes across the analyzed datasets. The consistency was calculated by computing a consistency score as follows: $C_g = \frac{\text{abs}(\text{upDatasets} - \text{downDatasets})}{\text{abs}(\text{upDatasets} + \text{downDatasets})}$ being “upDatasets” the number of datasets in which the gene (g) is upregulated, while “downDatasets” the number of datasets in which the gene (g) is downregulated.

Integrated large-scale transcriptome analysis

We calculated the number of datasets in which every gene considered in the analysis resulted to be differentially expressed and used this information to order the genes from the most frequently to the least frequently altered. Next, we identified the pathways that were significantly overrepresented among the lists of differentially expressed genes in each dataset. Similarly, to the gene-wise evaluation, we computed the frequency of significant alteration of each pathway across the datasets. Gene ID conversions were performed through the use of the *bioMart* [23] and the *clusterProfiler* [24] Bioconductor packages. The functional annotation was performed by using the *ReactomePA* R package [25]. We defined a pathway to be dysregulated when the overrepresentation test FDR adjusted p value ≤ 0.05 .

The gene and pathway rankings were carried out through the use of custom R scripts (<https://github.com/antoniofederico87/PSOnet>).

Data scaling

All of the collected microarray datasets were combined for cross-platform normalization. In particular, the *pamr* R package (version 1.56.1) [26] was used to mean-adjust the combined microarray data based on a batch variable representing the different datasets downloaded from

GEO. The outcome of the data scaling in mitigating the batch effect is shown in Additional file 1: Fig. S1.

Integrated Psoriasis Knowledge Base construction

We built a comprehensive gene-centric annotation, namely Integrated Psoriasis Knowledge Base (IPKB), reporting aggregated information about psoriasis from several categories of databases. In detail, the IPKB contains information annotated in 14 databases, grouped in 6 categories: druggability/tractability, genetic association, cell line-specific expression profiles, HumanKO/Trial, immune pathways and modules, and literature-derived PSO association, for a total of 22 gene sets (Additional file 1: Fig. S2). Column specifications of the IPKB are reported in Additional file 1: Table S2, while its visual organization is shown in Additional file 1: Fig. S5.

The IPKB was constructed by collecting data from numerous publications and/or public databases. Available psoriasis genetic data were retrieved from the NHGRI-EBI GWAS catalog of published genome-wide association studies [27] by using the keywords “*Psoriasis*” and “*Psoriasis vulgaris*” and selecting the genes with association p value $\leq 1E-05$, and the Open Targets database [28], selecting genes with genetic association score ≥ 0.1 for further analyses. Small molecule and antibody tractability data were also retrieved from Open Targets. Small molecule and biologics druggability data were collected from Finan et al. [29]. Protein localization data were downloaded from Pharos [30], Human Protein Atlas [31] (URL: <http://www.proteinatlas.org>) and from Uva et al. [32]. Immune pathway modules were retrieved from the Reactome database [33]. Human knockout (KO) data were from Saleheen et al. [34] and from Narasimhan et al. [35]. Immune cell-specific scRNA-Seq transcriptional signatures were collected from the Human Protein Atlas. The IPKB is publicly available in Zenodo (<https://doi.org/10.5281/zenodo.4740406>).

Networks inference and analysis

Two distinct co-expression networks were inferred by using the gene expression profiles of the lesional and non-lesional skin samples over all the included studies and the genes common to all the platforms. The co-expression networks were inferred through the use of the *INFORM* algorithm [36]. We set up *INFORM* in order to build a robust consensus network by using the *clr* [37], *aracne* [38] and *mrnet* [39] algorithms with the following correlation and mutual information measures: Pearson correlation, Kendall correlation, Spearman correlation, empirical mutual information, Miller–Madow asymptotic bias-corrected empirical estimator, Schurmann–Grassberger estimate of the entropy of a Dirichlet probability distribution and a shrinkage estimate of the

entropy of a Dirichlet probability distribution, as implemented in the *minet* Bioconductor package [40]. In order to carry out a network community detection, we used the Walktrap algorithm [41], implemented in *INfORM*. All computations performed on the inferred networks were carried out through the use of the *igraph* Bioconductor package [42].

Functional annotation

The functional annotations carried out in this study were based on the Reactome biological pathways and performed through the use of the *ReactomePA* [25] and *clusterProfiler* Bioconductor packages [24]. Moreover, the STRING database [43] was used to inspect the functional characteristics of the bridge genes.

Visualization

Visualization of the results was performed through the use of the *ggplot2* [44] and *gplots* [45] Bioconductor packages. The rendering of co-expression networks was performed by employing the *Gephi* software [46]. In this manuscript, we show a reduced representation of the actual networks in order to facilitate visualization.

Differential centrality analysis

For each of the networks, their node betweenness, closeness and degree centralities were calculated with the Python's *NetworkX* package (Python 3.6, NetworkX 2.3). The nodes were ranked according to each of the centrality measures. For each of the networks, their nodes' median rank based on the rankings of the three centrality measures were calculated. To compare the network of the lesional skin with the non-lesional one, the absolute difference between the median ranks of the two networks was calculated and the genes were ranked accordingly.

Gene set enrichment analysis

One tail gene set enrichment analyses (GSEA) were performed through Kolmogorov–Smirnov statistics, as implemented in the *stats* R package. Overrepresentation tests were performed by using the *bc3net* CRAN package [47].

Druggability evaluation of the lesional network

The druggability evaluation of the PSO lesional network was performed by using the DrugBank annotation (version 5.1.7) [48]. The Anatomical Therapeutic Chemical (ATC) Classification System was retrieved from the *joseitung/atc* GitHub R package. In order to increase the specificity of our analysis, we retrieved the drug–target associations from DrugBank and considered only drugs whose targets are included in one module. The analysis

was performed by considering the level 2 of the ATC codes annotation.

Results

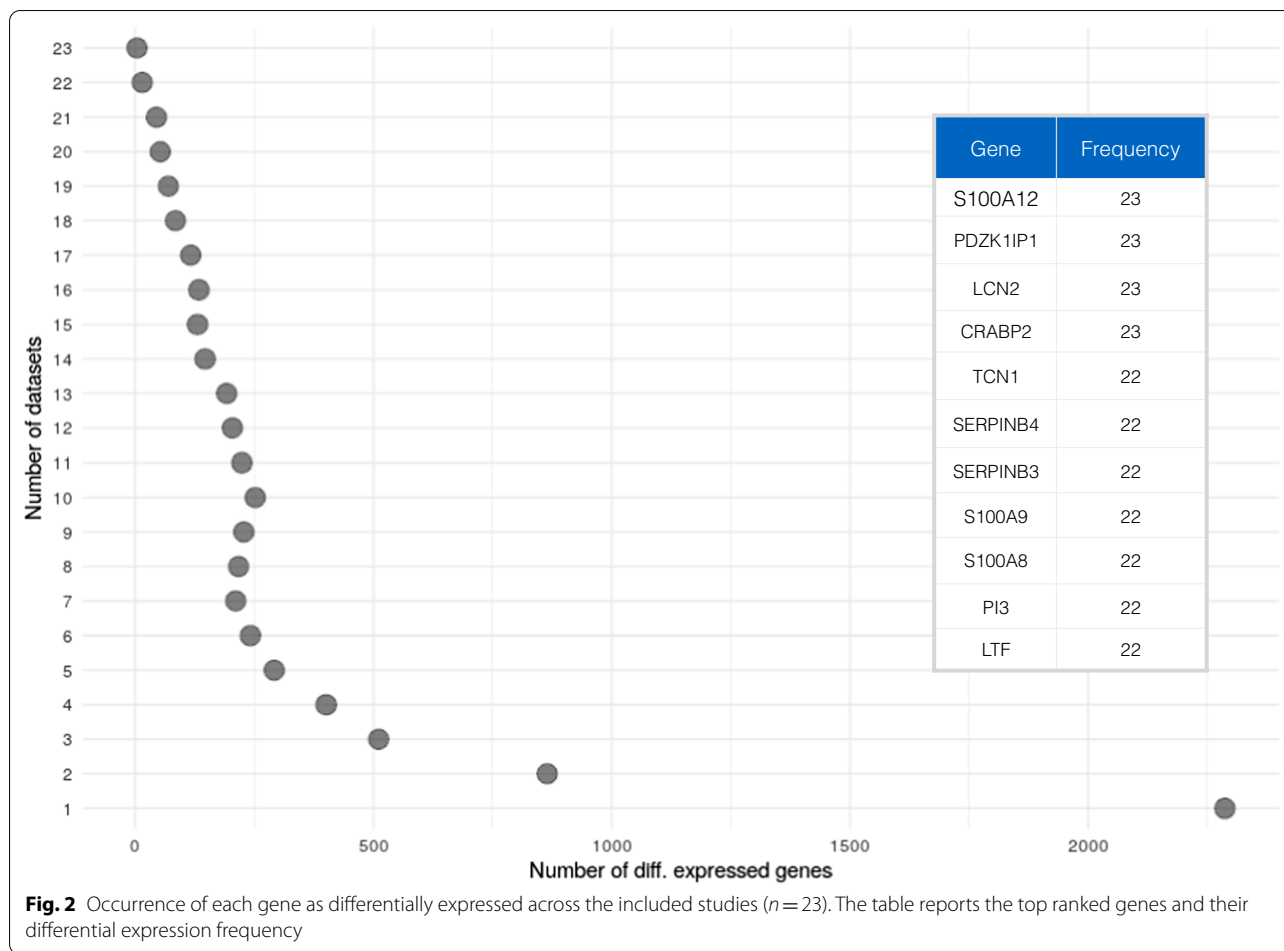
S100A12, *PDZK1IP1*, *LCN2* and *CRABP2* are the most commonly upregulated genes in the psoriatic lesion

In order to identify genes that are consistently dysregulated in transcriptomic studies of lesional skin samples with respect to non-lesional counterparts, we first analyzed each dataset individually. The number of differentially expressed genes in each dataset ranged from 3717 in GSE67853 to 100 in GSE57376, with a median of 1863 (Additional file 1: Fig. S2). The limited amount of clinical data did not allow us to infer any relationship between the number of differentially expressed genes and phenotypic characteristics of the patients. In this regard, we verified whether exists a linear relationship between the number of differentially expressed genes and the sample size of the analyzed datasets. Our analysis showed that the sample size has no impact on the amount of differentially expressed genes identified in each dataset (Additional file 1: Fig. S3). Therefore, we ranked the differential expressed genes on the basis of their occurrence across all the datasets. As a result, *S100A12*, *PDZK1IP1*, *LCN2*, and *CRABP2* genes were found to be differentially expressed in all 23 PSO datasets. Additional genes belonging to the *S100* and *Serpina* transcription factor families were differentially expressed in 22 out of 23 datasets. Overall, 92 genes were differentially expressed in at least 20 datasets. The top 100 ranked gene list derived from the integrated gene expression analysis is reported in Additional file 1: Table S3.

We then assessed which genes showed the highest magnitude of deregulation across all the datasets. Therefore, we ranked each differentially expressed gene in each dataset by a significance score, calculated as follows:

$$ss = -\log(FC) \cdot \log(\text{adj}pval)$$

where *FC* is the fold change between the mean of the expression of the lesional samples and the mean of the expression of the non-lesional counterparts; *adjpval* is the Benjamini–Hochberg [49] adjusted *p* value as obtained from the differential expression analysis. Our analysis highlighted *SERPINA4*, *S100A12*, and *TCN1* as the most dysregulated genes over all the datasets (Fig. 2). Among the frequently upregulated genes, *SERPINA4* showed a median logFC across the datasets of 6.3 [Q1: 5.4; Q3: 7.2] with a maximum of 7.8 in GSE13355; *S100A12* showed a median logFC of 5.0 [Q1: 4.5; Q3: 6.0] and a maximum of 6.7 in GSE30768; and *TCN1* had a median value of 5.1 [Q1: 4.1; Q3: 5.4] and a maximum of 7.2 in GSE57376. On the other hand, the top genes found to be downregulated in most of the datasets were *BTC*,



with a median logFC of -3.0 [Q1: -3.3 ; Q3: -2.6], and the strongest downregulation reported in GSE50790; *WIF1*, with a median logFC of -2.5 [Q1: -2.7 ; Q3: -2.3] and a maximum downregulation in GSE50790; and *PM20D1* with a median of 2.6 [Q1: -2.9 ; Q3: -2.0] and the lowest logFC of -4.5 in GSE47751. Moreover, we assessed the consistency of deregulation of differentially expressed genes across the analyzed datasets. The consistency was measured by taking into account the number of datasets in which genes showed the same trend of deregulation (up- or downregulated, see Methods). As a result, 98.7% of the genes that resulted to be differentially expressed in at least 2 datasets showed a full consistency of deregulation across the datasets (Additional file 1: Fig. S4).

Network analysis highlights genes with aberrant co-expression patterns

The integrated gene expression analysis allowed us to identify genes that are dysregulated in the psoriatic lesion with respect to the non-lesional skin, giving a quantitative perspective of the molecular alterations at

a transcriptional level of the disease. However, the integrated expression analysis uncovers only one aspect of the deregulation underlying the psoriatic phenotype. In fact, the molecular buildup of a tissue is not only determined by the expression patterns of individual genes, but also by their co-expression relationships. Therefore, to characterize the complex landscape of transcriptional alterations that sustain psoriasis, we identified disrupted patterns of gene co-expression. To do so, we inferred two transcriptome-wide gene co-expression networks from both the lesional and the non-lesional skin sample sets, respectively.

Since the networks were built from all the genes common to all of the microarray platforms, both of the networks are composed of 7310 genes, while the lesional network has 1,136,431 edges and the non-lesional one has 1559,790 edges.

The patterns of molecular alterations underlying the psoriatic phenotype can be characterized by investigating intrinsic topological properties of the inferred networks. One aspect that defines the differences between the two networks (lesional and non-lesional) is the centrality

of their genes, a property measuring the number of co-expression connections that a certain gene holds with the others.

The differential centrality analysis allowed us to identify genes with a significant difference of connectivity between the lesional and non-lesional network. Table 1 shows genes whose connectivity is heavily affected by psoriasis, since they are highly central in the lesional network while their centrality is lower in the non-lesional network. Specifically, the connectivity of *SERPINB4*, *KYNU*, *S100A12*, *CASP5*, *CXCL1*, *CXCL8* and *PNP* is the most affected. By comparing the results of the differential centrality analysis with the gene rank obtained by the integrated gene expression analysis, we observed that several of the top differentially central genes (DCG) were also differentially expressed in a large number of datasets. For instance, *SERPINB4*, *KYNU*, *S100A12*, *PNP* and *CXCL1*, which are among the top 10 DCG, resulted to be

differentially expressed in more than 20 datasets. On the contrary, some genes such as *YPEL1* and *HUS1* appear at the top of the DCG but are not differentially expressed in any of the collected datasets.

On the other hand, we identified a second set of DCG, which showed an opposite pattern of aberrant connectivity compared to the genes reported in Table 1. Indeed, the connectivity of a number of genes is affected so that the genes are highly central in the non-lesional network while they show a lower centrality in the lesional one (Table 2). Therefore, these genes lose a high number of co-expression connections in the psoriatic lesion in respect of the uninvolved skin. *IHH*, *AQP9*, *ITGB8*, *CD55*, *CMA1* are the most affected ones, showing this trend of connectivity. Interestingly, their frequency of differential expression in the integrated expression analysis is markedly low, being detected as differentially expressed in a maximum of 2 datasets, with the exception of *AQP9*, detected

Table 1 Top ten differentially central genes (DCG) between the lesional and the non-lesional network

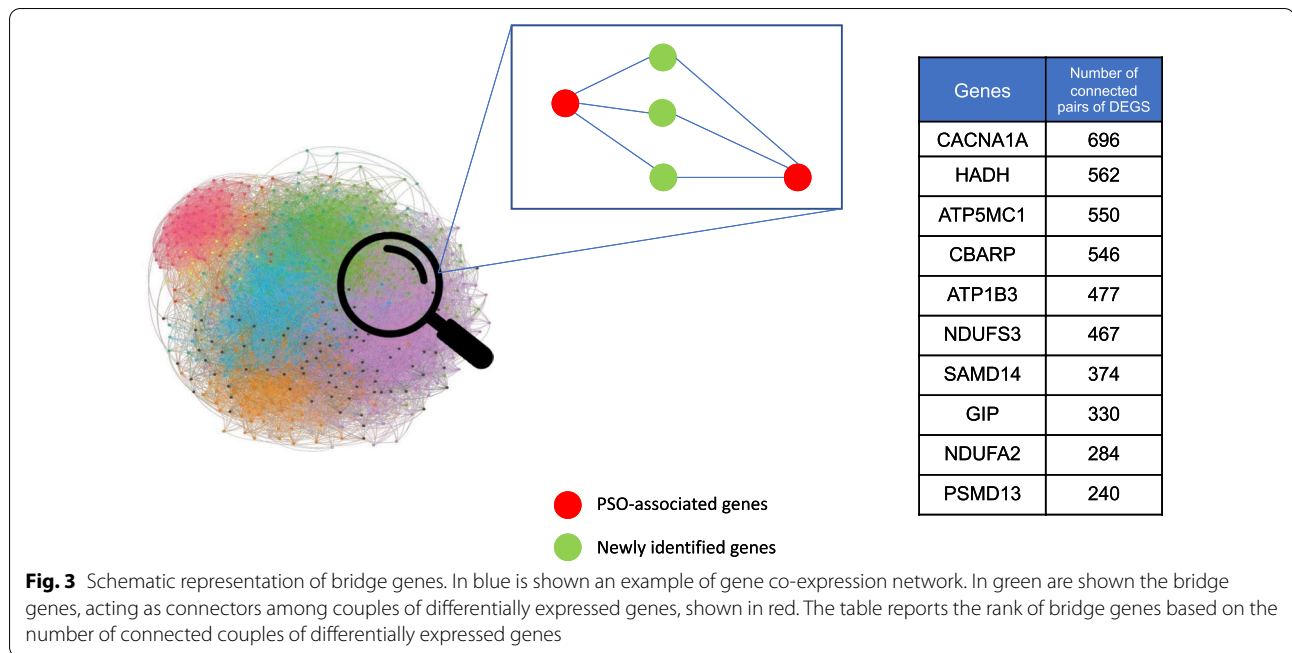
Top differentially central genes between lesional and non-lesional networks	Rank position in the lesional network	Rank position in the non-lesional network	Difference between the networks' ranks	Frequency of differential expression in the integrated expression analysis
SERPINB4	552	7128	6576	22
KYNU	911	7046	6135	22
S100A12	697	6620	5923	23
CASP5	1043	6875	5832	11
CXCL1	1107	6218	5783	20
CXCL8	1781	7135	5354	18
SLC23A2	1637	6983	5346	21
YPEL1	1336	6675	5339	0
PNP	1177	6377	5200	21
HUS1	1048	6232	5184	0

The rank position in the lesional and non-lesional network, the difference between the ranks and the frequency of differential expression of each gene are reported

Table 2 Top ten differentially central genes (DCG) between the lesional and the non-lesional network

Top differentially central genes between lesional and non-lesional networks	Rank position in the lesional network	Rank position in the non-lesional network	Difference between the networks' ranks	Frequency of differential expression in the integrated expression analysis
IHH	6575	1447	5128	0
AQP9	5756	631	5125	17
ITGB8	6880	1780	5100	1
CD55	6023	949	5074	1
CMA1	5874	852	5022	2
SRPK3	6711	1693	5018	0
ITPR1	6458	1529	4929	1
TUBGCP3	5865	971	4894	0
SYCP2	6570	1820	4750	0
NMB	6389	1726	4663	3

The rank position in the lesional and non-lesional network, the difference between the ranks and the frequency of differential expression of each gene are reported



in 17 studies. An overview of the impact of PSO on the co-expression connections in both the lesional and non-lesional network is shown in Additional file 1: Fig. S6.

Identification of novel candidate genes associated with psoriasis

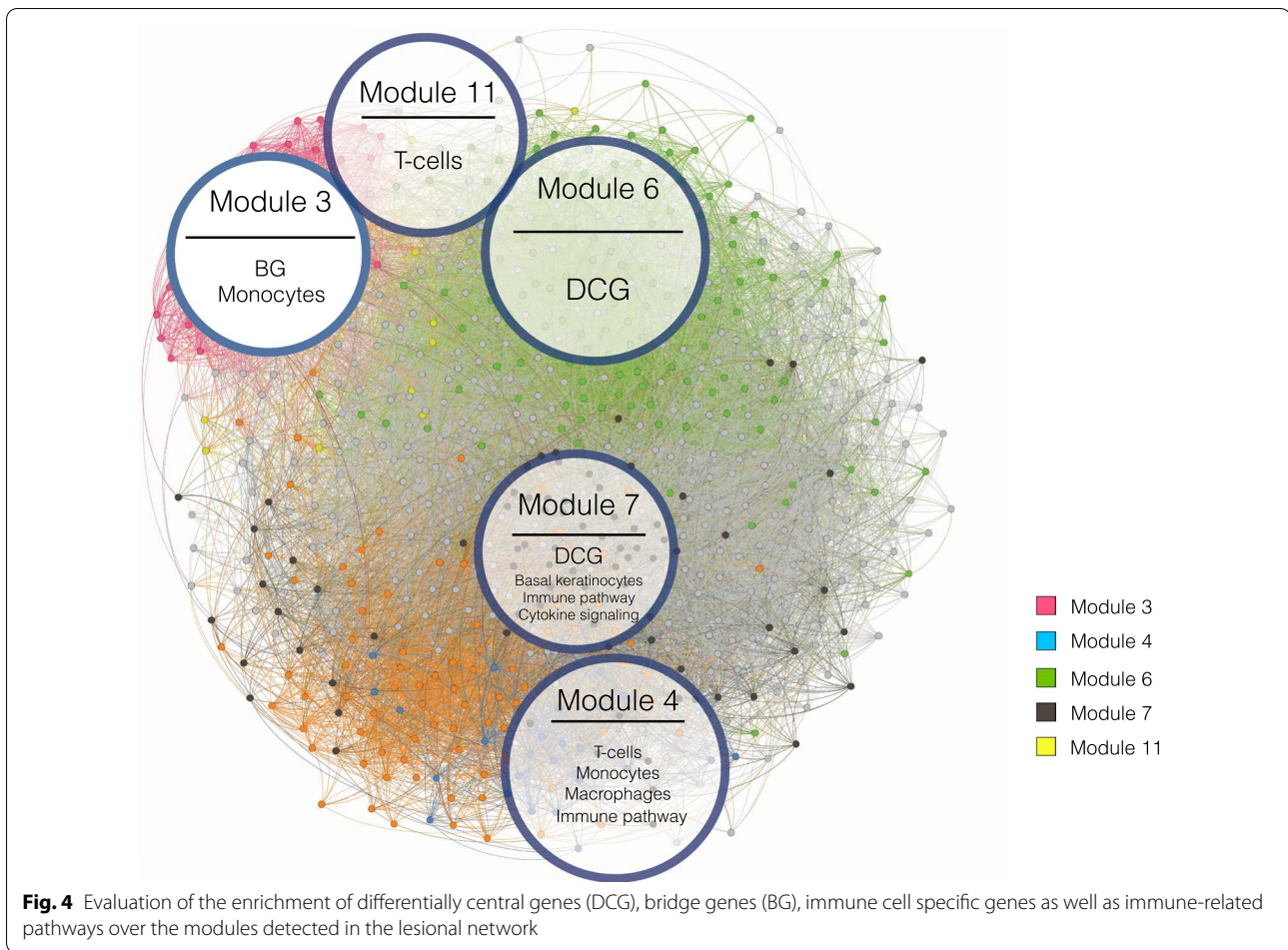
We hypothesized that, by studying the connectivity patterns among known psoriasis genes, it is possible to identify additional associated genes. Hence, a gene that is connected to two or more known psoriasis-associated genes is a strong candidate to be involved in its pathogenesis (Fig. 3). Based on this principle, we identified all the genes connecting pairs of differentially expressed genes (previously identified by the integrated gene expression analysis) within each of the networks (lesional and non-lesional, respectively), and hence acting as a bridge (hereafter referred to as “bridge genes”). By this analysis, we obtained a set of 1622 and 1940 bridge genes (BG) for the lesional network and non-lesional networks, respectively. Consequently, we selected a set of 250 genes acting as bridges in the lesional network, but not in the non-lesional one (Fig. 3). Among the bridge genes connecting a large number of differentially expressed gene pairs, we identified *CACNA1A* (Calcium Voltage-Gated Channel Subunit Alpha1 A) and its negative regulator *CBARP*, connecting 696 and 562 gene pairs, respectively. Likewise, the genes *HADH* and *ATP5MC1*, whose protein products function in mitochondria, connect a high number of dysregulated gene pairs (562 and 550, respectively) in the lesional network.

In order to characterize the functional properties of the bridge genes, we performed a functional annotation by using the STRING database. STRING shows a clustering of gene products involved in *RNA splicing*, which is the first enriched term in the gene ontology (GO) biological process, followed by cellular *nitrogen compound metabolic process* (both with FDR=0.0063) (Additional file 1: Fig. S7).

Network analysis allows the identification of disease-relevant communities

It is a widespread assumption that genes which are tightly co-expressed (whose expression levels are highly correlated) are likely to be also co-regulated, as well as involved in common functions [50]. Graph models allow the identification and characterization of such communities of genes. In this study, we investigated the arrangement of co-expressed genes in both the lesional and non-lesional networks by performing a community detection analysis. Thus, we identified 13 communities of co-expressed genes in the lesional network. The biggest community encompasses 1,888 genes, while the smallest 1 gene, with a median size of 309. In the non-lesional network, we identified 10 communities with a median size of 756 genes, with the biggest composed of 1,723 and the smallest by 1 gene. All our analyses were limited to modules composed of at least 10 genes (Additional file 1: Figs. S8 and S9).

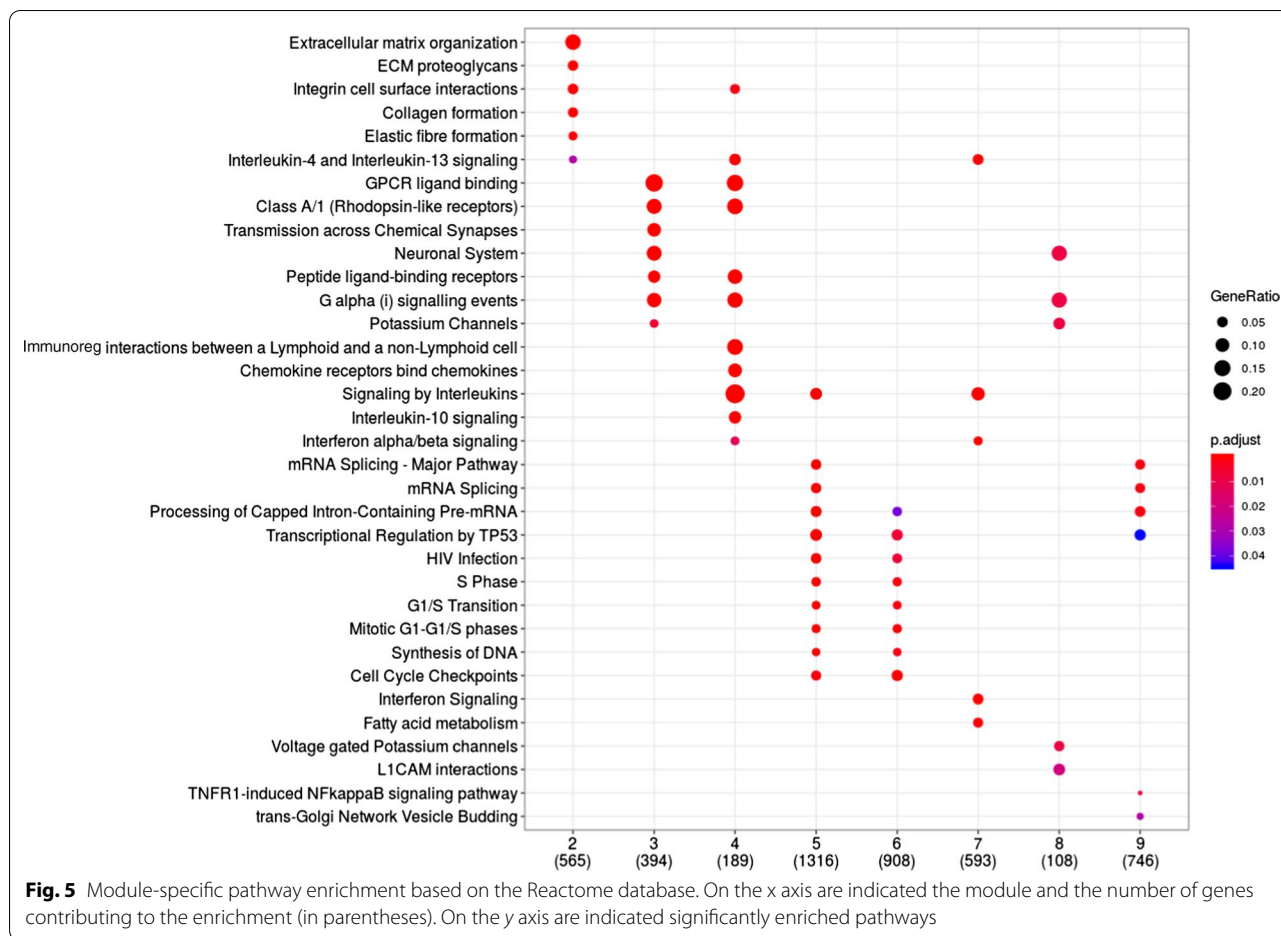
An interesting aspect we investigated is whether one or more network communities exist that enrich the



putative psoriasis-associated genes identified through the integrated gene expression analysis. In order to fulfill this aim, we performed a GSEA on the gene rank derived from the integrated gene expression analysis over the identified modules of the lesional network. As a result, we obtained that module 2, module 4 and module 7 significantly enriched the genes at the top of the integrated gene expression analysis rank ($p=2.41e-23$, $p=1.25e-24$, $p=4.89e-07$, respectively). Similarly, we performed the same analysis to assess whether the communities of the lesional network enrich for genes whose centrality is significantly different between the lesional and the non-lesional network, previously identified by the differential centrality analysis. We observed that module 6 and module 7 significantly enrich the top differentially central genes ($p=0.0054$ and $p=0.0026$, respectively). Finally, the same analysis was performed on the bridge genes set, in order to verify their enrichment over the modules. We found that module 3 significantly enriches the bridge genes ($p=1.57e-07$) (Fig. 4).

We characterized the biological functions of the gene communities identified in the lesional network ($fdr < 0.05$,

Fig. 5). Module 2, which is overrepresented by top-ranked genes of the integrated gene expression analysis, is significantly enriched in genes belonging to the extracellular matrix organization and proteoglycans, *collagen formation*, *integrin cell surface interactions* among others, which are expected in skin diseases like psoriasis. *Interleukin 4 and interleukin 13 signaling* pathways are also enriched by module 2 and 7. Moreover, *interleukins signaling* pathway is also enriched in module 7, together with other immunological pathways, such as *interferon signaling* pathway. However, Module 4 shows the strongest immunological signature among all. In fact, the genes belonging to this module significantly enrich *interleukin 10 signaling*, *interleukin 4 and 10 signaling* and *interferon alpha/beta signaling*. Modules 5 and 6 overrepresent pathways related to generic cell cycle functions, like *G1/S transition*, *S phase*, *transcriptional regulation of P53*, *mRNA splicing*. Likewise, module 3, which is overrepresented by bridge genes, enriches mostly for receptorial functions, such as *G-Protein Coupled Receptors (GPCRs) ligand binding*, *rhodopsin-like receptors*, and *peptide ligand-binding receptors*.



Prior knowledge enables the characterization of lesional network gene modules

We further performed an overrepresentation analysis of the gene sets collected in the IPKB in each community of the lesional network. This analysis highlighted that, for the category Human KO/Trial, Module 4 is significantly enriched in two out of three gene sets, HumanKOPakistan and ClinicalTrial ($p=0.027$ and $p=0.001$, respectively). Module 3 and 7 enrich for genes belonging to the HumanKOPakistan and HumanKOBritishPakistani sets ($p=0.003$ and $p=0.001$, respectively). Module 4 is also enriched for almost all the gene sets of the category Immune Pathways/Modules, with the most significant $p=5.06E-18$ in ImmunePathwayAdaptive. On the other hand, Module 7 enriches the ImmunePathwayCytokineSignalling ($p=2.12E-05$).

Additionally, since psoriasis poses its roots in the impairment of the immuno-inflammatory homeostasis, we wondered whether the modules of the lesional network are enriched by genes expressed in a specific manner in immune cell lines, which are primarily involved

in the aberrant response in psoriasis. By exploiting publicly available immune cell type-specific gene expression signatures from the Human Protein Atlas database, which we included in the DGI, we performed a GSEA to assess the enrichment of cell type-specific genes over the modules of the lesional network. We obtained that module 4 is enriched by genes specifically expressed in T cells ($p=0.006$), monocytes ($p=7.93e-05$) and macrophages ($p=0.015$). Similarly, module 3 is enriched by monocyte-associated genes ($p=0.035$) and module 7 by genes specifically expressed in basal keratinocytes ($p=0.01$) (Fig. 4).

Immunomodulators and dermatological drugs target specific modules of the lesional skin network

We characterized the druggability potential of the relevant modules identified in the previous analytical steps. To this end, we defined module-specific drug-target gene maps by exploiting publicly available information available at DrugBank. All of the modules except module 5 and 9 encompass a number of drugs which is higher than the number of genes composing the module (Fig. 6).

Moreover, by considering modules with a number of genes >10 and taking into account the frequency of drug-gene interactions over the total number of genes composing each module, we found that modules 3, 8, 7, 2 and 4 show a higher amount of drug target genes (38%, 38%, 31%, 31% and 30%, respectively) compared to other modules.

To further characterize the drugs mapping onto the lesional network, we restricted our analysis to drugs whose targets belong to one specific module only. By applying this restraint, we carried out further analyses on 3090 out of 5317 drugs.

We characterized the module-specific drugs on the basis of their therapeutic class as annotated in the second level of the World Health Organization (WHO) Anatomical Therapeutic Chemical (ATC) classification system. Module 8 contains the highest number of drugs mapped with respect to the size of the module (140 drugs, drugs/genes ratio=0.92), followed by module 11 (87 drugs, drugs/genes ratio=0.68), module 2 (500 drugs, drugs/genes ratio=0.61), module 3 (354 drugs, drugs/genes ratio=0.61) and module 6 (729 drugs, drugs/genes ratio=0.59). Module 4, which we previously identified to have a marked immunological profile, encompasses target genes for 106 drugs, showing a drugs/genes ratio of 0.41.

While the most represented drug category in module 8 is anti-emetics and anti-nauseants (A04), for both module 8 and module 11 dermatologicals belonging to the

anti-acne preparations category (D10) are predominantly represented (Fig. 7). Potassium Voltage-Gated Channel Subfamily H Member 2 (*KCNH2*) and the Retinoic Acid Receptor Alpha (*RARA*) play a pivotal role in the drug-gene interactions of module 8. In fact, *KCNH2* protein is a target of a high number of drugs, including Erythromycin and Chlorbutanol, while the retinoic acid receptor alpha is targeted by dermatological compounds including tretinoin, isotretinoin and adapalene. Interestingly, *RARA* is also targeted by two other retinoids employed in the treatment of severe psoriasis, Tazarotene and Etretinate. In module 11, Dapsone and Resorcinol target the *NAT2* and *TPO* gene products, respectively. In module 4, the most represented class of compounds is immunosuppressant (L04). In fact, Alefacept, targeting the T cell surface antigen CD2, together with Abatacept and Belatacept, targeting the T cell activation antigen CD86 are among the numerous molecules belonging to this category. Also in module 4 Framycetin is represented in the medicated dressings category (D09), which is known to act on the *CXCR4* gene product. Interestingly, module 4 also encompasses a number of target genes for immunostimulant molecules (L03). Our analysis highlighted that *IL2RA* and *IL2RB* are targets of Aldesleukin, a compound employed in *IL2* replacement therapies, while the Colony Stimulating Factor 3 Receptor (*CSF3R*) is targeted by several immunostimulant drugs, such as Filgrastim, Lenograstim, Pegfilgrastim and Lipegfilgrastim. Finally,

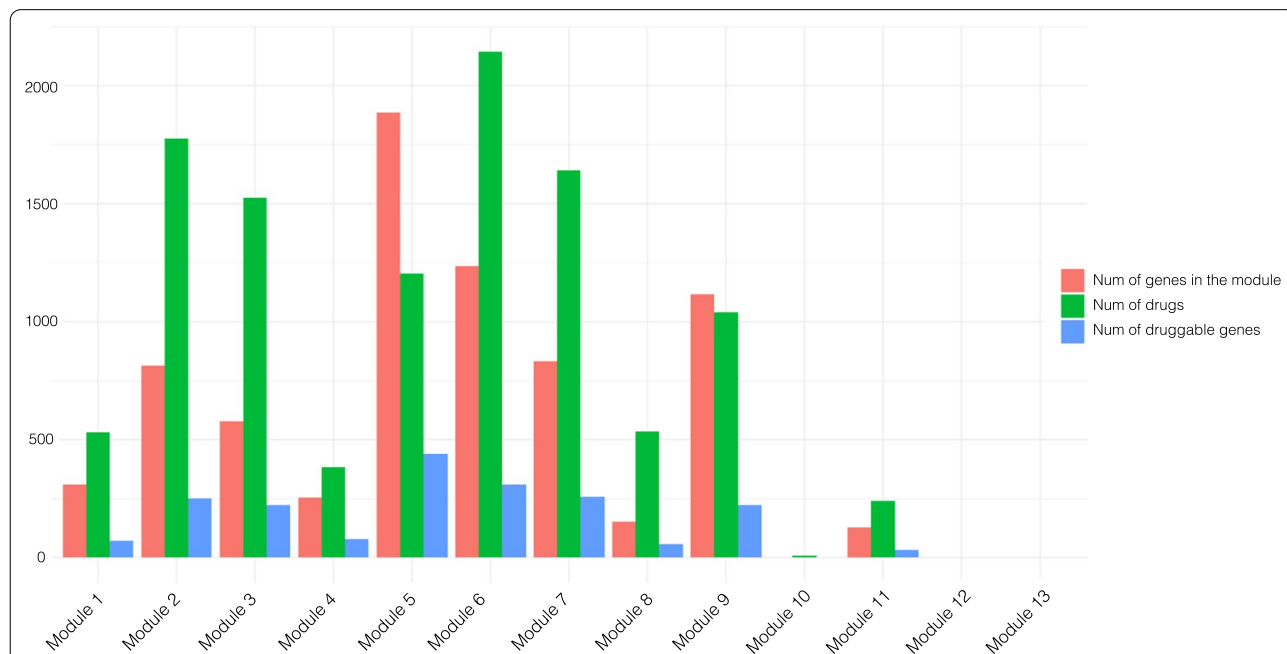
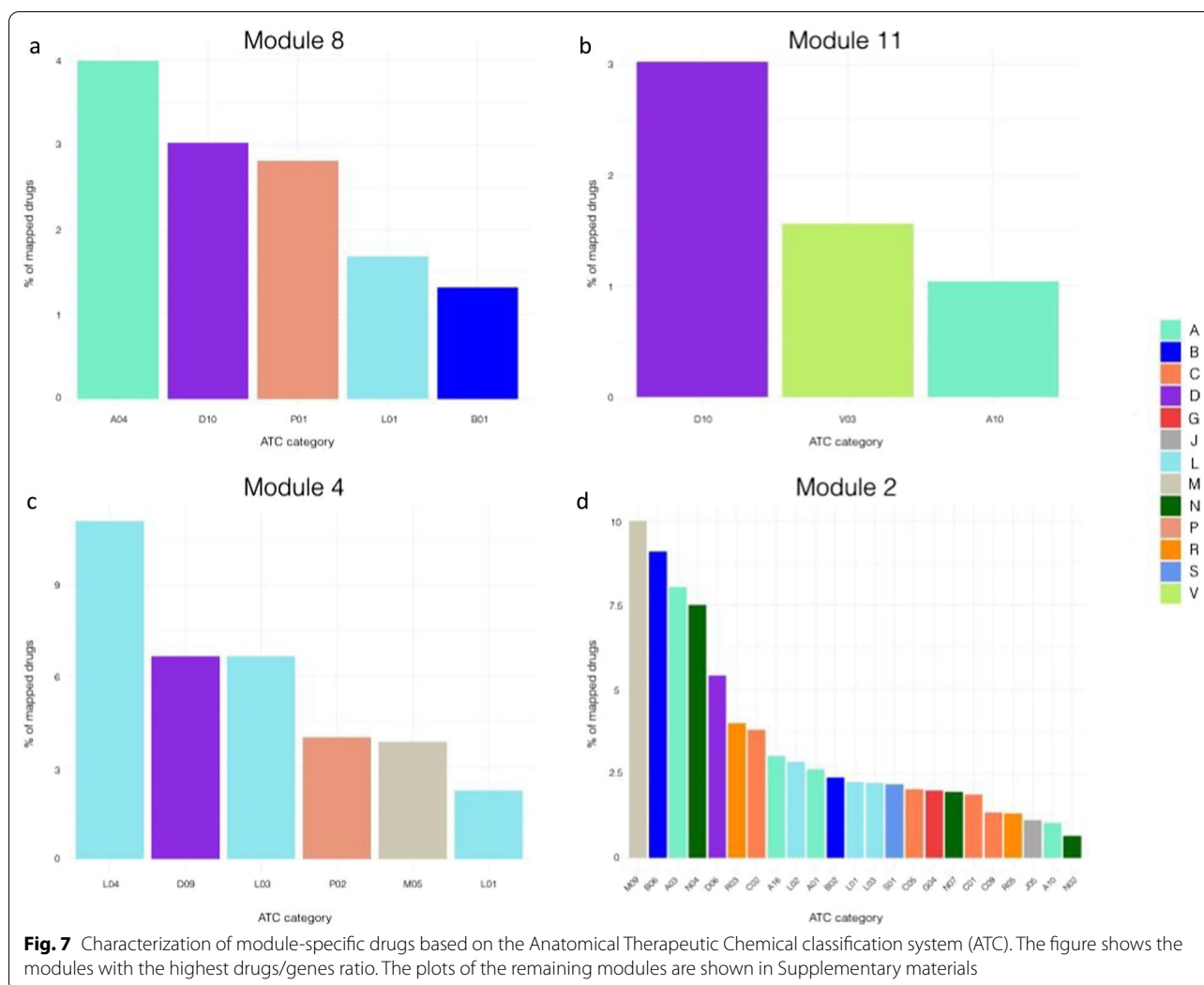


Fig. 6 Evaluation of the module-based druggability profile in the PSO lesional network. In light red is shown the number of genes composing the module, in green the number of drugs, and in blue the number of druggable genes



module 2 is enriched by a wide spectrum of pharmacological categories, ranging from drugs employed in the treatment of musculoskeletal disorders (M09), hematological agents (B06) to drugs used for gastrointestinal disorders (A03) and anti-Parkinson drugs (N06). Given the possibility that the druggability signature observed in the lesional network model might be influenced by ongoing or past therapeutic treatments at the moment of the sampling rather than pathophysiological mechanisms of psoriasis, and the lack of clinical information does not allow to establish the magnitude of this effect, we repeated the analysis on both the lesional and non-lesional networks by focusing only on drug targets that resulted consistently differentially expressed across the analyzed datasets (Additional file 1: Fig. S10). Common patterns of druggability retrieved in both lesional and non-lesional networks might result either from systemic treatments or from pathological alterations of the molecular buildup of

the skin of psoriatic patients. On the other hand, druggability profiles that are specific only to the lesional network should capture potential molecular mechanisms that can be used to specifically target localized alterations at the level of the lesions. Therefore, we first computed the module similarities among the two networks based on the gene content (Additional file 1: Fig. S10, panel A). We found a significant similarity between module 2 of the lesional network (L) and the module 3 of the non-lesional network (NL), module 4 (L)-module 7 (NL), and module 8 (L)-module 1 (NL). Subsequently, we profiled the druggability of such modules (Additional file 1: Fig. S10, panel B). Interestingly, we found anti-acne preparations (D10) to be specific of the lesional network in module 8 and being the second most represented category. Similarly, module 2 of the lesional network shows a specificity for hematological agents (B06) and antibiotics and chemotherapeutics for dermatological use (D06).

Discussion

In this study, we analyzed a large collection of transcriptomics datasets recently curated [17] in order to gain new knowledge about complex patterns of gene alteration with a role in the psoriatic phenotype.

By analyzing 23 transcriptomics datasets, we identified genes and pathways that are consistently dysregulated in the psoriatic lesion as compared to uninvolved skin. Our analysis found *CRABP2*, *LCN2*, *S100A12* and *PDZ-KIP1* dysregulated in all the datasets, suggesting their importance in the definition of the psoriatic phenotype. Overall, our differential expression analysis highlighted the upregulation of genes involved in inflammatory cascades, such as *S100A12*, *SERPINB4* and *TCN1*, and the pronounced downregulation of genes related to developmental pathways, such as epidermal growth factor family members (*BTC*) and genes involved in WNT signaling (*WIFI1*). These results are in line with findings previously reported by Swindell and colleagues in 2013, where they compared gene expression from integrating different transcriptomics datasets [51]. In particular, the authors showed that the upregulation of *S100A12*, *SERPINB4* and *TCN1* is specifically marked in keratinocytes as compared to other cell types. Moreover, genes most strongly decreased in psoriatic skin (such as *BTC* and *WIFI1*), most were weakly expressed in myeloid-derived cell types, but did show specific expression in epidermis.

In addition to known genes associated with psoriasis, we also prioritized novel candidates, such as *SYNCRIP* encoding a protein involved in the control of translation such as alternative splicing and mRNA maturation [52], as well as *SASH3* whose protein product could function as a signaling adapter protein in lymphocytes [53].

Since it is well known that genes act in a coordinated manner in both physiological and pathological conditions, we inferred and analyzed co-expression network models representative of the psoriatic lesion and the uninvolved skin in order to identify the disrupted patterns of gene co-expression underlying the psoriatic lesion.

In the context of graph models, genes are co-expressed with variable numbers of other genes (interactors), signifying their relative importance in defining the phenotype underlying the gene network. We identified the genes with the most different number of interactors in the two networks derived from lesional and non-lesional samples, respectively, highlighting two important aspects. First, the deregulation of distinct genes in the lesional skin affects the co-expression relationships with other genes, which are not necessarily dysregulated. Second, this highlights the importance of going beyond the classical gene expression analysis, which is focused on the

evaluation of individual genes, failing to capture the complex relationships sustaining biological processes. In fact, we identified a few genes with an aberrant co-expression connectivity in the lesional network as compared with the non-lesional one, which do not show a differential expression, such as *YPEL1* and *HUS1*. While *YPEL1* may play a role in the regulation of cell division and in the polarization of fibroblasts toward an epithelial-like morphology [54], *HUS1* product is involved in cell cycle arrest following DNA damage. In fact, *HUS1* gene is one of the top associated genes with Xeroderma pigmentosum, since it is responsible for the impaired repair capacity of UV-mediated DNA damages.

While differential expression analysis identified genes dysregulated in psoriatic lesions, gene network analysis investigating the connectivity patterns among such genes in the lesional network revealed 250 non-dysregulated genes connecting (bridging) dysregulated ones. As previous attempts to identify genes associated with psoriasis by transcriptomics relied mainly on differential expression, it is not surprising that 223 of our 250 newly identified genes have not been associated with psoriasis so far according to Open Targets. Thus, we here describe a completely new group of genes related to transcriptional deregulation in psoriatic lesions, which we call “bridge genes,” since they connect couples of differentially expressed genes within the lesional network, and, therefore, they are putatively associated with psoriasis.

The bridge gene connecting the highest number of differentially expressed genes is *CACNA1A* (Calcium Voltage-Gated Channel Subunit Alpha1 A). This gene encodes a calcium channel, which regulates intracellular processes such as contraction, secretion, neurotransmission and gene expression, suggesting that bridge genes have superior/broad-spectrum roles in cell regulation. *CACNA1A* is not only followed by its related gene *CBARP* (CACN Subunit Beta Associated Regulatory Protein), but also by a number of genes involved in mitochondrial metabolic activities such as *HADH*, whose enzymatic activity is exploited in the fatty acid beta-oxidation process, and *ATP5MC1*, coding for a subunit of mitochondrial ATP synthase and is responsible for the synthesis of ATP during oxidative phosphorylation by exploiting the protonic gradient across the mitochondrial inner membrane.

Interestingly, we found that many bridge genes are significantly co-expressed within module 3, which is enriched by genes involved in biological processes such as GPCR ligand binding, transmission across chemical synapses, and potassium channels indicating that bridge genes are related to broadly receptorial functions.

Pilar Pedro et al. [55] reports about the role of the GPCRs in the translation of extracellular signals into intracellular cascades that regulate the activation of keratinocytes proliferation and differentiation, including major signaling pathways, such as Hedgehog, Hippo YAP1 and WNT/B-catenin. In the same work, the authors underline the role of the neural–epithelial connection, mediated by β -adrenergic receptor (β AR) signaling in triggering keratinocyte proliferation, which is over-activated in the psoriatic lesion.

Moreover, module 3 is particularly rich in genes associated with monocytes. The role of hyper-reactive monocytes in the psoriatic phenotype has long been known [56]. In fact, Golden et al. observed elevated adhesion of monocytes and, in turn, increased formation of aggregates, which they also correlated with disease severity and underlying a major role for innate immunity in disease progression [57].

Module 4 is largely characterized by dysregulated genes whose activity lies in immune-related pathways, including signaling by interleukins, interferon signaling and chemokines and their receptors. Moreover, it encompasses genes which are specifically expressed in a reservoir of immune cell types, such as T cells, monocytes and macrophages, underlying its role in the chronic auto-inflammatory response characteristic of psoriasis. Indeed, a pivotal role for T cells and cells of the myeloid lineage, including monocytes and macrophages, is well established [58–63].

The immune-related nature of module 4 is reflected also by the druggability analysis of the lesional network model. In fact, several genes belonging to module 4 are targets of both immunostimulant and immunosuppressive drugs, such as interleukins and chemokines. This suggests that this module could be a good reservoir of putatively novel pharmacological targets for the development of therapeutic approaches with an immunomodulatory action to treat psoriatic lesions. Along with these categories of compounds, dermatological medications were also represented in module 4. Framycetin (also known as neomycin sulfate), among others, is a neomycin component employed in the treatment of ocular and skin bacterial infections. To the best of our knowledge, this compound is currently not employed for the treatment of the psoriatic plaques. Furthermore, module 8 showed an interesting *scenario* regarding its drug target content. In fact, we found that the retinoic acid receptor alpha (*RARA*) is the target of a plethora of chemical compounds already employed in the treatment of severe psoriasis plaques. For instance, the topical agent Tazarotene and oral agent Acitretin (and its predecessor Etretinate) are compounds largely used in the treatment of psoriatic plaques [64, 65]. Tazarotene

is a retinoid drug which has been approved in 2019 by Food and Drug Administration (FDA) in combination with Halobetasol in psoriasis-affected adults [66]. On the other hand, Acitretin is used in severe psoriatic manifestations, but its high lipophilic capacity shows teratogenic effects and it is contraindicated in pregnancy for 3 years prior to conception [67]. Etretinate, a metabolic product of Acitretin, is a high lipophilic retinoid which was used in severe psoriatic manifestations [68], but its use was suspended between 1996 and 1998 for its teratogenic effects [69].

The limited amount of clinical data made available along with the transcriptional profiles annotated in public repositories poses some limitations to the present study. The lack of detailed clinical information makes the identification of gene markers or co-expression communities associated with clinical characteristics impossible, hampering the predictive power of the present study. A notable difference in the amount of differentially expressed genes arises from the included datasets. The lack of clinical information does not allow to infer any meaningful relationship between the number of dysregulated genes and certain clinical parameters of affected patients, including severity of the disease and elapsed time from the first diagnosis. This drawback poses, as a further limitation, the lack of information that could suggest whether one or multiple patients underwent active pharmacological therapy. Indeed, the transcriptional signatures underlying the topology of both the lesional and the non-lesional network models might be influenced by a possible (ongoing or terminated) pharmacological therapy to which some of the patients might be subjected, interfering with the pharmacological footprint that we investigated. Moreover, this hinders the possibility of translating this study to a precision medicine level, making possible the characterization of the impaired molecular relationships at a single-patient resolution.

In conclusion, in this study we combined an integrative gene expression analysis with co-expression network analysis in order to identify novel aspects of the psoriatic lesion at a molecular level. Our approach allowed us to give an insight into the known alterations associated with psoriasis by identifying novel genes which can putatively act as disease biomarkers. Future mechanistic studies elucidate their role in the disease onset and progression, while epidemiological studies will be necessary to assess their clinical relevance.

Abbreviations

ATC: Anatomical Therapeutic Chemical Classification System; GEO: Gene Expression Omnibus; GSEA: Gene set enrichment analysis; GWAS: Genome-wide association study; IPKB: Integrated Psoriasis Knowledge Base; PSO: Psoriasis.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40246-022-00431-x>.

Additional file 1. Additional material including supplementary figures and tables.

Acknowledgements

AF was supported by Tampere Institute for Advanced Studies (IAS).

Author contributions

AF and DG designed the study. AF performed the analyses. AP, LM and VF contributed to the analyses. DMK and EDR participated in data generation and curation. AF, LM, GDG, HN, EVDB, JR, SE, KE and DG interpreted the results. AF, LM, GDG, CS, SW and DG wrote the manuscript. DG supervised the study. All authors read and approved the final manuscript.

Funding

This study was supported by the EU IMI2 Biomap Project (Grant agreement 821511), Academy of Finland (Grant agreement 322761) and the European Research Council (ERC) programme, Consolidator project “ARCHIMEDES” (Grant agreement 101043848).

Availability of data and materials

The data used in the present study are publicly available in Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>). The GEO IDs of the utilized datasets are reported in Additional file 1: Table S1. A priori data have been included in the Integrated Psoriasis Knowledgebase (IPKB), and it is publicly available in Zenodo (<https://doi.org/10.5281/zenodo.4740406>). All codes written and used for this study are available at: <https://github.com/antoniofederico87/PSOnet>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Faculty of Medicine and Health Technology, Tampere University, Kauppi Campus, Arvo Ylpön Katu 34, 33520 Tampere, Finland. ²BioMeditech Institute, Tampere University, Tampere, Finland. ³Finnish Hub for Development and Validation of Integrated Approaches (FHAIVE), Tampere University, Tampere, Finland. ⁴Tampere Institute for Advanced Studies, Tampere University, Tampere, Finland. ⁵Sanofi Immunology and Inflammation Research Therapeutic Area, Precision Immunology Cluster, Cambridge, Massachusetts, USA. ⁶Institute of Biomedicine, University of Eastern Finland, Kuopio, Finland. ⁷Department of Dermatology, Radboud University Medical Center, Radboud Institute for Molecular Life Sciences, Nijmegen, The Netherlands. ⁸Immunology Therapeutic Area, UCB Pharma, Slough, UK. ⁹Department of Dermatology and Allergy, Technical University of Munich, Munich, Germany. ¹⁰Unit of Dermatology and Venerology, Department of Medicine, Karolinska Institute, Karolinska University Hospital, Stockholm, Sweden. ¹¹ZAUM-Center of Allergy and Environment, Technical University and Helmholtz Center Munich, Munich, Germany. ¹²St. John's Institute of Dermatology, King's College London, London, UK. ¹³Department of Dermatology, Kiel University, 24105 Kiel, Germany. ¹⁴Institute of Biotechnology, University of Helsinki, Helsinki, Finland.

Received: 5 September 2022 Accepted: 7 November 2022

Published online: 28 November 2022

References

- Albanesi C, Madonna S, Gisondi P, Girolomoni G. The interplay between keratinocytes and immune cells in the pathogenesis of psoriasis. *Front Immunol.* 2018;9:1549.
- Dand N, Mahil SK, Capon F, Smith CH, Simpson MA, Barker JN. Psoriasis and genetics. *Acta Derm Venereol.* 2020;100:adv00030.
- Ayala-Fontánez N, Soler DC, McCormick TS. Current knowledge on psoriasis and autoimmune diseases. *Psoriasis (Auckl).* 2016;6:7–32.
- Tsoi LC, Rodriguez E, Degenhardt F, Baurecht H, Wehkamp U, Volks N, et al. Atopic dermatitis is an IL-13-dominant disease with greater molecular heterogeneity compared to psoriasis. *J Invest Dermatol.* 2019;139:1480–9.
- Kulski JK, Kenworthy W, Bellgard M, Taplin R, Okamoto K, Oka A, et al. Gene expression profiling of Japanese psoriatic skin reveals an increased activity in molecular stress and immune response signals. *J Mol Med.* 2005;83:964–75.
- Reischl J, Schwenke S, Beekman JM, Mrowietz U, Stürzebecher S, Heubach JF. Increased expression of Wnt5a in psoriatic plaques. *J Invest Dermatol.* 2007;127:163–9.
- Yao Y, Richman L, Morehouse C, de los Reyes M, Higgs BW, Boutrín A, et al. Type I interferon: potential therapeutic target for psoriasis? *PLoS ONE.* 2008;3:e2737.
- Suárez-Fariñas M, Li K, Fuentes-Duculan J, Hayden K, Brodmerkel C, Krueger JG. Expanding the psoriasis disease profile: interrogation of the skin and serum of patients with moderate-to-severe psoriasis. *J Invest Dermatol.* 2012;132:2552–64.
- Piruzian E, Bruskin S, Ishkin A, Abdeev R, Moshkovskii S, Melnik S, et al. Integrated network analysis of transcriptomic and proteomic data in psoriasis. *BMC Syst Biol.* 2010;4:41.
- Piruzian ES, Sobolev VV, Abdeev RM, Zolotareno AD, Nikolaev AA, Sarkisova MK, et al. Study of molecular mechanisms involved in the pathogenesis of immune-mediated inflammatory diseases, using psoriasis as a model. *Acta Naturae.* 2009;1:125–35.
- Pavel A, del Giudice G, Federico A, Di Lieto A, Kinaret PAS, Serra A, et al. Integrated network analysis reveals new genes suggesting COVID-19 chronic effects and treatment. *Brief Bioinform.* 2021.
- Sheils TK, Mathias SL, Kelleher KJ, Siramshetty VB, Nguyen D-T, Bologa CG, et al. TCRD and pharos 2021: mining the human proteome for disease biology. *Nucleic Acids Res.* 2021;49:D1334–46.
- Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, et al. Impact of genetic polymorphisms on human immune cell gene expression. *Cell.* 2018;175:1701–15.
- Barabási A-L, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. *Nat Rev Genet.* 2011;12:56–68.
- van Dam S, Vösa U, van der Graaf A, Franke L, de Magalhães JP. Gene co-expression analysis for functional classification and gene-disease predictions. *Brief Bioinform.* 2018;19:575–92.
- Ahn R, Gupta R, Lai K, Chopra N, Arron ST, Liao W. Network analysis of psoriasis reveals biological pathways and roles for coding and long non-coding RNAs. *BMC Genomics.* 2016;17:841.
- Federico A, Hautanen V, Christian N, Kremer A, Serra A, Greco D. Manually curated and harmonised transcriptomics datasets of psoriasis and atopic dermatitis patients. *Sci Data.* 2020;7:343.
- Marwah VS, Scala G, Kinaret PAS, Serra A, Alenius H, Fortino V, et al. eUTOPIA: solUTION for omics data preprocessing and analysis. *Source Code Biol Med.* 2019;14:1.
- Bolstad BM, Collin F, Simpson KM, Irizarry RA, Speed TP. Experimental design and low-level analysis of microarray data. *Int Rev Neurobiol.* 2004;60:25–58.
- Fasold M, Binder H. AffyRNAdegradation: control and correction of RNA quality effects in GeneChip expression data. *Bioinformatics.* 2013;29:129–31.
- 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.
- Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics.* 2012;28:882–3.

23. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc.* 2009;4:1184–91.
24. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16:284–7.
25. Yu G, He Q-Y. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol Biosyst.* 2016;12:477–9.
26. Hastie T, Tibshirani R, Narasimhan B, Chu G. pamr: Pam: prediction analysis for microarrays. 2019.
27. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* 2019;47:D1005–12.
28. Carvalho-Silva D, Pierleoni A, Pignatelli M, Ong C, Fumis L, Karamanis N, et al. Open targets platform: new developments and updates two years on. *Nucleic Acids Res.* 2019;47:D1056–65.
29. Finan C, Gaulton A, Kruger FA, Lumbers RT, Shah T, Engmann J, et al. The druggable genome and support for target identification and validation in drug development. *Sci Transl Med.* 2017;9.
30. Nguyen D-T, Mathias S, Bologna C, Brunak S, Fernandez N, Gaulton A, et al. Pharos: collating protein information to shed light on the druggable genome. *Nucleic Acids Res.* 2017;45:D995–1002.
31. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, et al. Tissue-based map of the human proteome. *Science.* 2015;347:1260419.
32. Uva P, Lahm A, Sbardellati A, Grigoriadis A, Tutt A, de Rinaldis E. Comparative Membranome expression analysis in primary tumors and derived cell lines. *PLoS ONE.* 2010;5:e11742.
33. 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.
34. Saleheen D, Natarajan P, Armean IM, Zhao W, Rasheed A, Khetarpal SA, et al. Human knockouts and phenotypic analysis in a cohort with a high rate of consanguinity. *Nature.* 2017;544:235–9.
35. Narasimhan VM, Hunt KA, Mason D, Baker CL, Karczewski KJ, Barnes MR, et al. Health and population effects of rare gene knockouts in adult humans with related parents. *Science.* 2016;352:474–7.
36. Marwah VS, Kinaret PAS, Serra A, Scala G, Lauerma A, Fortino V, et al. Inform: inference of network response modules. *Bioinformatics.* 2018;34:2136–8.
37. Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, Cottarel G, et al. Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol.* 2007;5:e8.
38. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Dalla Favera R, et al. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinform.* 2006;7(Suppl 1):S7.
39. Meyer PE, Kontos K, Lafitte F, Bontempi G. Information-theoretic inference of large transcriptional regulatory networks. *EURASIP J Bioinform Syst Biol.* 2007;7:9879.
40. Meyer PE, Lafitte F, Bontempi G. minet: A R/Bioconductor package for inferring large transcriptional networks using mutual information. *BMC Bioinform.* 2008;9:461.
41. Pons P, Latapy M. Computing communities in large networks using random walks. In: Yolum plnar, Güngör T, Gürgeç F, Özturan C, editors. *Computer and Information Sciences - ISICIS 2005.* Berlin, Heidelberg: Springer, Berlin; 2005. pp. 284–93.
42. Csardi G, Nepusz T. The igraph software package for complex network research. *Int J Complex Syst.* 2006;1695:1–9.
43. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019;47:D607–13.
44. Wickham H. ggplot2 - Elegant graphics for data analysis. New York: Springer; 2016.
45. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Wolfgang Huber Andy Liaw, Lumley T, et al. gplots: Various R programming tools for plotting data; 2016.
46. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. Unpublished. 2009.
47. de Matos SR, Emmert-Streib F. Bagging statistical network inference from large-scale gene expression data. *PLoS ONE.* 2012;7:e33624.
48. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* 2018;46:D1074–82.
49. Controlling the false discovery rate: a practical and powerful approach to multiple testing on JSTOR. <https://www.jstor.org/stable/2346101>. Accessed 15 Apr 2021.
50. Barabási A-L, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet.* 2004;5:101–13.
51. 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.
52. Mizutani A, Fukuda M, Ibata K, Shiraishi Y, Mikoshiba K. SYNCRIP, a cytoplasmic counterpart of heterogeneous nuclear ribonucleoprotein R, interacts with ubiquitous synaptotagmin isoforms. *J Biol Chem.* 2000;275:9823–31.
53. Kukuk L, Dingley AJ, Granzin J, Nagel-Steger L, Thiagarajan-Rosenkranz P, Ciupka D, et al. Structure of the SLY1 SAM homodimer reveals a new interface for SAM domain self-association. *Sci Rep.* 2019;9:54.
54. Farlie P, Reid C, Wilcox S, Peeters J, Reed G, Newgreen D. Ypel1: a novel nuclear protein that induces an epithelial-like morphology in fibroblasts. *Genes Cells.* 2001;6:619–29.
55. Pedro MP, Lund K, Iglesias-Bartolome R. The landscape of GPCR signaling in the regulation of epidermal stem cell fate and skin homeostasis. *Stem Cells;* 2020.
56. Bar-Eli M, Gallily R, Cohen HA, Wahba A. Monocyte function in psoriasis. *J Invest Dermatol.* 1979;73:147–9.
57. Golden JB, Groft SG, Squeri MV, Debanne SM, Ward NL, McCormick TS, et al. Chronic psoriatic skin inflammation leads to increased monocyte adhesion and aggregation. *J Immunol.* 2015;195:2006–18.
58. Greb JE, Goldminz AM, Elder JT, Lebwohl MG, Gladman DD, Wu JJ, et al. Psoriasis. *Nat Rev Dis Primers.* 2016;2:16082.
59. Kim J, Krueger JG. The immunopathogenesis of psoriasis. *Dermatol Clin.* 2015;33:13–23.
60. Wang H, Peters T, Kess D, Sindrilaru A, Oreshkova T, Van Rooijen N, et al. Activated macrophages are essential in a murine model for T cell-mediated chronic psoriasisform skin inflammation. *J Clin Invest.* 2006;116:2105–14.
61. Stratis A, Pasparakis M, Rupec RA, Markur D, Hartmann K, Scharffetter-Kochanek K, et al. Pathogenic role for skin macrophages in a mouse model of keratinocyte-induced psoriasis-like skin inflammation. *J Clin Invest.* 2006;116:2094–104.
62. Leite Dantas R, Masemann D, Schied T, Bergmeier V, Vogl T, Loser K, et al. Macrophage-mediated psoriasis can be suppressed by regulatory T lymphocytes. *J Pathol.* 2016;240:366–77.
63. Ward NL, Loyd CM, Wolfram JA, Diaconu D, Michaels CM, McCormick TS. Depletion of antigen-presenting cells by clodronate liposomes reverses the psoriatic skin phenotype in KC-Tie2 mice. *Br J Dermatol.* 2011;164:750–8.
64. Tanghetti E, Lebwohl M, Stein GL. Tazarotene revisited: safety and efficacy in plaque psoriasis and its emerging role in treatment strategy. *J Drugs Dermatol.* 2018;17:1280–7.
65. Heath MS, Sahni DR, Curry ZA, Feldman SR. Pharmacokinetics of tazarotene and acitretin in psoriasis. *Expert Opin Drug Metab Toxicol.* 2018;14:919–27.
66. Radonjic A, Evans EL. A novel halobetasol propionate 0.01%/tazarotene 0.045% fixed combination treatment for psoriasis. *Dermatol Ther.* 2019;32:e12979.
67. Hoffman MB, Farhangian M, Feldman SR. Psoriasis during pregnancy: characteristics and important management recommendations. *Expert Rev Clin Immunol.* 2015;11:709–20.
68. Raposo I, Torres T. Palmoplantar psoriasis and palmoplantar pustulosis: current treatment and future prospects. *Am J Clin Dermatol.* 2016;17:349–58.
69. Geiger JM, Baudin M, Saurat JH. Teratogenic risk with etretinate and acitretin treatment. *Dermatology (Basel).* 1994;189:109–16.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.