Evaluation of the Psoriasis Transcriptome across Different Studies by Gene Set Enrichment Analysis (GSEA)

Mayte Suárez-Fariñas1, Michelle A. Lowes1, Lisa C. Zaba1, James G. Krueger1*

1 Laboratory for Investigative Dermatology, The Rockefeller University, New York, New York, United States of America, 2 Center for Clinical and Translational Science, The Rockefeller University, New York, New York, United States of America

Abstract

**Background:** Our objective was to develop a consistent molecular definition of psoriasis. There have been several published microarray studies of psoriasis, and we compared disease-related genes identified across these different studies of psoriasis with our own in order to establish a consensus.

**Methodology/Principal Findings:** We present a psoriasis transcriptome from a group of 15 patients enrolled in a clinical study, and assessed its biological validity using a set of important pathways known to be involved in psoriasis. We also identified a key set of cytokines that are now strongly implicated in driving disease-related pathology, but which are not detected well on gene array platforms and require more sensitive methods to measure mRNA levels in skin tissues. Comparison of our transcriptome with three other published lists of psoriasis genes showed apparent inconsistencies based on the number of overlapping genes. We extended the well-established approach of Gene Set Enrichment Analysis (GSEA) to compare a new study with these other published list of differentially expressed genes (DEG) in a more comprehensive manner. We applied our method to these three published psoriasis transcriptomes and found them to be in good agreement with our study.

**Conclusions/Significance:** Due to wide variability in clinical protocols, platform and sample handling, and subtle disease-related signals, intersection of published DEG lists was unable to establish consensus between studies. In order to leverage the power of multiple transcriptomes reported by several laboratories using different patients and protocols, more sophisticated methods like the extension of GSEA presented here, should be used in order to overcome the shortcomings of overlapping individual DEG approach.

Introduction

The study of human diseases such as psoriasis has benefited significantly from analysis of the transcriptome, the global gene expression of a diseased tissue compared to its healthy counterpart. However, as more studies are carried out independently in multiple laboratories, effective methodology to leverage these multiple studies becomes necessary. Such methodologies have significant hurdles to overcome: first, multiple studies are likely to use different platforms, different sample dissection, handling and preparation, and, especially, different definition of the non-diseased counterpart, resulting in different physical samples being hybridized against different platforms [1,2,3]. Second, computational analysis and statistical treatment required to assess the transcriptome are just as likely to be considerably different.

In many instances, all that is available from published studies are lists of differentially expressed genes (DEG). It is tempting to evaluate the agreement between studies simply by evaluating the intersection between the published lists, the “Venn diagram approach”. However, such an approach suffers serious methodological shortcomings [4,5,6]. Use of the original raw data of the studies has shown that studies which are apparently discordant in terms of their overlapping individual DEG lists are, in fact, both concordant and predictive [4,5]. However, most of the time the original raw data is unavailable, and furthermore a complete reanalysis of all data is needlessly laborious. In such cases use of the published lists of DEG is a necessity. Here we present an extension to the widely used Gene Set Enrichment Analysis (GSEA) method, where it suffices to have full access to the complete list of gene expression values for a single study, while the remaining studies only require the DEG list.

In the last few years, the use of Gene-Set approach had emerged as a powerful tool to identify sets of functionally related genes or pathways that are associated with a disease phenotype [7,8]. Gene-Set based methods were designed to address limitations of conventional single gene methods [6] by evaluating
differential expression patterns of gene groups instead of individual genes. GSEA, introduced by Mootha et al. [9] and further developed by Subramanian et al. [10], was one of the first methods using the Gene-Sets approach, and is arguably the most widely used of such methods. Here we use GSEA as a conventional approach to identify pathways related to the psoriatic phenotype. Furthermore, we propose to extend the use of GSEA as a tool to easily cross-compare prior lists of DEG genes.

We developed this method specifically to compare several high-quality studies that defined the psoriasis transcriptome by identifying DEG between psoriatic lesions and non-lesional tissue from the same patients [11,12,13,14,15]. Those studies had identified key genes involved in psoriasis pathogenesis, using a non-biased approach. Because the genomic data for more recent studies is more comprehensive than in the earlier studies due to the larger number of genes represented in the latest Affymetrix chips, we have compared our transcriptome with the three others described above, and introduce the concept of using GSEA as a more robust way of comparing genomic data.

Results
Disease-modulated genes

The analysis of our data identified a psoriasis transcriptome composed of 732 up-regulated probesets (representing 579 genes with unique ENTREZ identifier) and 890 down-regulated probesets (703 genes) with fold change (FCH) greater than 2 and false discovery rate (FDR) less than 0.05 (Table 1, and Table S1). Certain genes with low expression on the Affymetrix chip were confirmed by RT-PCR, and will be discussed in the next section.

To further consider the biological significance of our data, we used GSEA in the classical manner, to identify pathways that correlate with the psoriatic phenotype [10,18]. GSEA evaluates how genes in queried pathways are distributed in the fold change (lesional versus non-lesional) order list generated by our data (all probesets included). This is quantified by using the Enrichment Score (ES), a weighted Kolmogorov-Smirnov-like statistic that evaluates if the members of the pathway are randomly distributed or found at the extremes (top or bottom) of the list. If genes in a pathway rank at the top of the new fold change list, i.e., they are overrepresented at the top, then the enrichment score (ES) will be close to 1. Conversely if the ES = −1, then genes are overrepresented at the bottom of our fold change data. A perfect agreement is reached if ES = 1 for the up-regulated genes and ES = −1 for the down-regulated genes. A normalized enrichment score (NES) takes into account the number of genes in the pathway. A positive NES indicates that the list of genes is enriched at the ‘‘top’’ of the ordered fold change list, and a negative NES indicates that the list in question is enriched at the ‘‘bottom’’ of the list.

GSEA may be used with well known ‘‘canonical’’ pathways and Gene ontology categories, but also with sets that contain genes sharing the same transcription factor binding site, the same microRNA binding motif or the same cis-regulatory motif. It can also be used with curated collections such as GeneSigDb (http://compbio.dfci.harvard.edu/gensigdb), which contain gene signatures of cancer, viral and stem cell biology, and the CGP collection of the MSigDb (http://www.broadinstitute.org/gsea/msigdb/) that contains gene expression signatures of genetic and chemical perturbations, or computational derived sets such as cancer modules (presented in [19]).

We queried our psoriasis transcriptome with a set of cytokine-treated keratinocyte pathways for IL-17, TNF, IL-17+TNF, IFNγ, and IL-22, reported in [17,20] and IL-1α [21]. GSEA showed that those pathways were enriched in lesional skin of psoriasis patients (Table 2). We used a list (“IL-17 Gaffen”) considered to be IL-17-gene targets defined by Shen et al [22]. This was also significantly enriched in our psoriasis transcriptome. An IFNα-keratinocyte pathway [14] was also significantly enriched in lesional skin, supporting the potential role for IFNα discussed by Yao et al [14].

A list of genes representing the transcriptome of inflammatory myeloid DCs [23] was also significantly enriched in lesional skin. In addition, the cell cycle and TLR signaling from the collection of canonical pathways (C2 CP) available at the Molecular Signature Database (MSigDb) were enriched in psoriasis lesional skin. Gudjonsson et al reported lack of evidence of enrichment of the Hedghog Signalling Pathways in psoriasis [24]. GSEA analysis did not detect any significant enrichment of this pathway with the psoriasis phenotype. (ES = 0.37, NES = 1.05 and p = 0.39).

Comparison with other published studies

We next performed a comparative analysis of the DEG lists of these three studies [12,14,15] with our transcriptome. Table 1 summarizes the main characteristics of the four studies. Zhou’s study which uses early hgu95 (a,b,c,d,e) chips, reported 397 up-regulated and 613 down-regulated probesets representing 270 and 397 unique genes respectively [15]. Yao et al conducted a study on

Table 1. Description of studies.

| Platform/chips | Zhou   | Yao                  | Gudjonsson | Suárez-Fariñas |
|----------------|--------|----------------------|------------|----------------|
| Sample size    | 16     | 26                   | 58         | 15             |
| Expression Algorithm | dChip  | GCRMA   | RMA       | GCRMA           |
| Statistical test | t-test | Sam (paired) | t-test | Paired t-test  |
| Multiple Hypothesis correction | none | FDR through permutations | FDR through permutations | FDR Benjamini-Hochberg |
| Cut-off | FCH>2, p<0.05 | FCH>2, q-value<0.05 | FCH>2, FDR<0.05 | FCH>2, FDR<0.05 |
| # of Up-regulated probesets (genes) | 397 ps (270 genes) | 1408 ps (974 genes) | 721 ps (508 genes) | 732 ps (579 genes) |
| # of Down-regulated probesets (genes) | 613 ps (397 genes) | 1465 ps (853 genes) | 364 ps (248 genes) | 890 ps (703 genes) |

doi:10.1371/journal.pone.0010247.t001
hgu133plus2 chips and reported 1408 up-regulated and 1465 down-regulated probesets (974 and 853 genes respectively) [14]. More recently, Gudjonsson et al. also using hgu133plus2 chips with a large sample size [12] reported a set of 721 up-regulated and 364 down-regulated probesets (508 genes and 248 genes).

Figure 1 shows a Venn diagram illustrating the intersection between the four studies. There were approximately 11,000 ENTREZ identifiers common to the 4 chip series, but only 126 genes were up-regulated (Figure 1A) and 38 down-regulated (Figure 1B) in all four studies. The numbers of upregulated genes in our transcriptome were similar to the Gudjonsson group, but less than Yao’s. The number of down-regulated genes were greater than the Gudjonsson group, and again, less than Yao’s.

The list of the DEG in common in the 4 studies (Table S2) contains many genes known to be upregulated in psoriasis, such as IFNγ-regulated genes STAT-1, STAT-3 and MX1; antimicrobial peptides (beta defensin 4, lipocalin 2, S100A7); and pro-inflammatory proteins such as IL-8, CXCL1, IL-1β, TNFSF10/TRAIL. However, some genes known to be upregulated in psoriasis were only identified in one study. For example STAT2 was only identified in our study, JAK3 only in Yao’s, and IL-12RB1 only in Zhou’s study. The list of genes that were down-regulated in all 4 studies include CCL27, also called Cutaneous T cell –attracting chemokine (CTACK), which has a role in memory T cell homing to the skin [25], and Aquaporin 9, a member of a family of proteins that form water channels across membranes [26].

Some well-recognized inflammatory genes involved in psoriasis were not detected by most of the four studies, for example IFNγ, IL-17, iNOS. This is due to the fact that the expression of these genes is usually low on the Affymetrix gene array platform (0–4 range of expression in log2-scale) and hence fold change is not accurately measured. Most analysis pipelines filter out low abundance genes so they may be excluded from the statistical analysis, or the resultant fold change is very low, albeit significant. This is a major limitation of the use of these arrays for the study of these genes.

To confirm the role of these genes in psoriasis, we analyzed RT-PCR data of many of these inflammatory genes. We used data from the same clinical trial of etanercept treatment of psoriasis [16], however, we compared only baseline lesional skin and non-lesional skin. The fold change and p-value for each gene by RT-PCR is presented in Table 3, and all of these genes except LTA and p35 were significantly different in lesional skin (p<0.05). We also present the fold change of any of these genes that were detectable in any of the microarray studies. It can be seen that IL-23p19, IL-12/23p40, IL-22, IFNγ, IL-6 were not found to be differentially expressed by any of the four studies, but the fold change by RT-PCR was greater than 4. Furthermore, IL-17, IL-20, CCL4, iNOS and CCL3 were detected in only one study, and with a lower fold change than detected by RT-PCR (which was greater than 5.5). We also included the percentage of samples in our study with low intensity, as defined as expression values less than 4. For example, p19 has expression of less than 4. It can be seen that the genes that were not detected by any or only one microarray study had low abundance in more than 87% of the samples.

Correlation of other published psoriasis transcriptomes compared to our transcriptome using Gene Set Enrichment Analysis (GSEA)

A more robust approach beyond a Venn diagram was required to overcome these limitations. We propose to use a Gene-Set approach to analyze how published DEG rank in our fold change data. This reduces the bias due to preprocessing steps, statistical protocols and stringency of cut-offs. This approach is successfully used in the Connectivity Map [27], a pattern recognition instance that correlates disease signatures (based on gene expression of any platform) with drugs. The idea is to use the GSEA framework [10] by considering the published DEG as a pathway or gene set, and quantify how well the up (and down) regulated genes rank in the ordered fold change for all genes in our data. This will generate an ES for up-regulated genes and one for the down-regulated genes. The connectivity score (CS) can be used to give a measure of agreement between studies, by combining the two ES into one final value, as used in the connectivity map to rank drugs that better correlate with disease. A value of CS near 1 would indicate perfect agreement between a study list and our analysis, whereas 0 would indicate no agreement, and −1 a negative correlation.

GSEA showed that there was highly significant enrichment of psoriasis DEGs, both up- and down-regulated genes, from the three studies compared to our data (Table 4). The GSEA plot for

| PATHWAYS | No. of genes in pathway | ES | NES | FDR |
|----------|-------------------------|----|-----|-----|
| IFNγ Up in KC (Yao) | 28 | 0.91 | 2.74 | <10⁻⁴ |
| IL-17 and TNF Up in KC | 30 | 0.89 | 2.69 | <10⁻⁴ |
| IL-17 Up in KC | 46 | 0.87 | 2.86 | <10⁻⁴ |
| IL-1 Up in KC | 34 | 0.85 | 2.66 | <10⁻⁴ |
| IL-17 GAFFEN | 27 | 0.81 | 2.43 | <10⁻⁴ |
| IFNγ Up in KC | 872 | 0.47 | 2.31 | <10⁻⁴ |
| TNF Up in KC | 472 | 0.46 | 2.14 | <10⁻⁴ |
| IL-22 Up in KC | 10 | 0.89 | 2.07 | <10⁻⁴ |
| Terminal Differentiation | 33 | 0.69 | 2.10 | <10⁻⁴ |
| Inflammatory myeloid DCs (psoriasis) | 121 | 0.42 | 1.68 | <10⁻⁴ |
| Cell Cycle (KEGG) | 64 | 0.58 | 2.06 | <10⁻⁴ |
| TLR signaling Pathway (KEGG) | 58 | 0.59 | 2.05 | <10⁻⁴ |
| Cytokine-Cytokine receptor interaction (KEGG) | 111 | 0.44 | 1.71 | 0.02 |

doi:10.1371/journal.pone.0010247.t002
up and down-regulated genes in Zhou’s, Yao’s and Gudjonsson’s lists is shown in Figure 2. P-values for ES and CS were calculated using 10,000 simulations. For Zhou’s list, the ES for the up-regulated genes was 0.86 ($p<0.0001$) and $-0.63$ ($p<0.0001$) for the down-regulated genes (Figure 2A). The CS = 0.75 ($p<0.0001$) indicates a positive significant agreement between Zhou’s signature and our data. For Yao’s transcriptome (Figure 2B), the ES = 0.90 ($p<0.0001$) for the up-regulated genes and ES = $-0.86$ ($p<0.0001$) for the down-regulated genes, for CS = 0.88 ($p<0.0001$), which also indicates a positive significant agreement. For Gudjonsson’s transcriptome (Figure 2C), the ES = 0.93 ($p<0.0001$) for the up-regulated genes and ES = $-0.91$ ($p<0.0001$) for the down-regulated genes, for CS = 0.92 ($p<0.0001$), which also indicates a positive significant agreement. In general, a better agreement was observed among up-regulated genes for all studies. In addition, Yao’s transcriptome correlated better with our study than Zhou’s, which is not surprising since the array series and the statistical protocols used in Yao’s and ours were more similar than those in Zhou’s.

We used the same approach to compare two published DEGs of other skin diseases produced by our group: squamous cell carcinoma (SCC) [28] and basal cell carcinoma (BCC) [29]. The CS for SCC was 0.69 and the CS of BCC was 0.42, considerably lower than in psoriasis (Table 4). This degree of enrichment is most likely reflects the origin of this data from our own lab, and the cutaneous nature of the specimens, as well as epidermal hyperproliferation and inflammation in all these three diseases.

Discussion

Investigators might be surprised at the lack of overlap between DEG lists, as shown in the Venn diagram (Figure 1). However, if one considers all the variables involved in the four studies, summarized in Table 1, it is not that surprising. Although all the studies were conducted using Affymetrix platform, they used different array series, which may contribute to variability in results. Besides the obvious laboratory effect due to sample preparation, technician experience, equipment calibration, and the use of different preprocessing algorithms [30], alternative statistical tests and stringent cut-offs also contribute to different results [3,31].

Measuring agreement of microarray studies by overlap of DEG lists generated by individual studies has been largely criticized [5] because it is highly inconsistent, even in the presence of small variation in the data as in the case of technical replicates [32,33]. A low overlap between DEG does not directly indicate low agreement between studies [3,31,33]. Here we present our new data on the psoriasis transcriptome from our patients, as well as a comparison of our data with three published DEG lists for psoriasis. We find only 164 genes in common for the four lists. However by changing the focus of the single-gene approach behind the intersection of DEG involved in psoriasis and using a gene set approach, a closer biological similarity between the studies is revealed. The extension of GSEA presented here enables us to see that cellular processes and molecular signature involved in psoriasis is very robust across the studies. In this paper, we extended the use of GSEA to compare new expression data with previously published DEG lists in order to validate psoriasis disease-related gene profiles. It is worth noting that this approach is applicable to expression data obtained through deep sequencing (potentially improving sensitivity for low abundance genes and cross-hybridization problems of current microarray technology). Moreover, this approach is easily extendable to other omics applications and more complex phenotypes. Since the method is based on ranking a list according to a phenotype, the ranked list can be derived from other measures besides gene expression fold changes from microarray chips or deep sequencing; such phenotype measures may include odd ratios of single nucleotide polymorphisms (SNPs), or a microRNA profile assay derived from microarray technologies or deep sequencing.

The use of GSEA as a gene set approach is not unique: extensions to GSEA [34] and other Gene Set methods and
statistics have also been proposed, and could also be used to compare transcriptomes. Efron and Tibshirani [35] proposed the MaxMean statistic instead of the weighted Kolmogorov Smirnov statistics used in the classical GSEA. Dinu et al [36] extended the single gene analysis SAM and proposed SAM-GS. See [7,37] for a comparative study of different gene set enrichment methods.

Extensions other than the classical difference between two phenotypes have also been reported. For example, we used the time slope of a mixed effect model as a phenotype to evaluate the time-response of cytokines pathways to psoriasis treatment with TNF inhibitor [17].

In this report, we show an excellent and simple method for researchers seeking validation of their own expression data with published lists from different studies.

Materials and Methods

Patients

Twenty adult patients with moderate to severe psoriasis were treated with etanercept 50 mg subcutaneously twice weekly for 12 weeks (clinical trial no. NCT00116181). The clinical and histological response of patients in this trial was previously published [16]. The gene array was performed on samples from 15 sequential patients [17]. RT-PCR was performed on samples from all 20 patients.

Ethics Statement

The clinical trial (no. NCT00116181) was conducted according to the principles expressed in the Declaration of Helsinki and informed consent for their information to be stored in the hospital database and used for research was obtained from all patients in written form. This research was conducted under protocol JKR-0542 approved by the Rockefeller University Institutional Review Board.

Table 3. RT-PCR validation.

| Gene Symbol | RT-PCR FCH (log2) | p-value | Microarray FCH (log2) | Suarez-Farinas | Zou | Yao | Gudjonsson | % Low Int |
|-------------|-------------------|---------|----------------------|----------------|-----|-----|------------|-----------|
| p19 IL23A   | 2.66              | 6.34    | 0.011                |                |     |     |            | 100       |
| p40 IL12B   | 4.03              | 16.32   | 1.8×10⁻⁶⁵           |                |     |     |            | 97        |
| LTA1        | 0.83              | 1.77    | 0.302                |                |     |     |            | 100       |
| IL22        | 3.96              | 15.53   | 1.5×10⁻⁴            |                |     |     |            | 100       |
| IFNγ        | 2.28              | 4.85    | 2.8×10⁻⁴            |                |     |     |            | 100       |
| IL4         | −1.45             | 0.36    | 0.034                |                |     |     |            | 100       |
| IL6         | 2.66              | 6.33    | 7.1×10⁻⁴            |                |     |     |            | 100       |
| IL17        | 6.17              | 71.87   | 3.8×10⁻⁴            |                |     |     |            | 100       |
| IL20        | 3.95              | 15.48   | 1.4×10⁻⁵            |                |     |     |            | 100       |
| CCL4        | 2.83              | 7.12    | 9.2×10⁻⁵            |                |     |     |            | 100       |
| IL22        | 5.45              | 43.57   | 1.7×10⁻⁹            |                |     |     |            | 100       |
| IL18        | 3.66              | 12.66   | 6.8×10⁻⁶            |                |     |     |            | 100       |
| IL19        | 5.10              | 34.34   | 8.3×10⁻⁹            |                |     |     |            | 100       |
| K16         | 3.56              | 11.83   | 4.1×10⁻⁵            |                |     |     |            | 100       |
| MX1         | 3.66              | 17.42   | 6.5×10⁻⁷            |                |     |     |            | 100       |
| IL8         | 6.19              | 72.82   | 1.0×10⁻⁷            |                |     |     |            | 100       |
| J Defensin  | 4.12              | 17.42   | 6.5×10⁻⁷            |                |     |     |            | 100       |

1The largest fold change (FCH) was reported when there were several probesets representing the same gene.
2Percentage of samples with low intensity as defined by having expression smaller that 4.

Table 4. GESA analysis of published transcriptomes with our data.

| Gene Set     | No. of genes in pathway | ES    | NES   | FDR   | CS   |
|--------------|-------------------------|-------|-------|-------|------|
| Gudjonsson - UP | 386                      | 0.93  | 4.25  | <10⁻⁴ | 0.92 |
| Gudjonsson - Down | 153                     | −0.91 | −3.77 | <10⁻⁴ |     |
| Yao - UP      | 670                      | 0.90  | 4.31  | <10⁻⁴ | 0.88 |
| Yao - Down    | 487                      | −0.86 | −4.06 | <10⁻⁴ |     |
| Zhou - UP     | 199                      | 0.86  | 3.69  | <10⁻⁴ | 0.75 |
| Zhou - Down   | 227                      | −0.63 | −2.71 | <10⁻⁴ |     |
| SCC LS vs NL UP | 859                     | 0.69  | 3.37  | <10⁻⁴ | 0.69 |
| SCC LS vs NL Down | 655                    | −0.70 | −3.41 | <10⁻⁴ |     |
| BCC LS vs Normal UP | 191                  | 0.38  | 1.61  | <10⁻⁴ | 0.42 |
| BCC LS vs Normal Down | 326                 | −0.46 | −2.09 | <10⁻⁴ |     |

ES: Enrichment Score; NES: Normalized Enrichment Score; FDR: False Discovery Rate.

doi:10.1371/journal.pone.0010247.t003

doi:10.1371/journal.pone.0010247.t004
Sample Preparation and Hybridization

RNA was extracted from skin biopsies and hybridized to Affymetrix hgu133a2 chips as described in [17].

RTPCR

Primers and Probes for TaqMan RT-PCR assays had been previously described [16]. New primers used in this study were CCL3 (Hs00234142_m1), CCL4 (Hs99999148_m1), AREG (Hs00950669_m1), IL-19 (Hs00604657_m1). All assays were obtained from Applied Biosystems. Data was normalized using human acidic ribosomal protein as a housekeeping gene [16].

Statistical Analysis

Expression values were obtained using gema algorithm. Samples were filtered for unreliable low expression value and low variability. To compare lesional with non-lesional values, the moderated t-test were filtered for unreliable low expression value and low variability. All analysis was carried out using R programming language (www.R-project.org) and Bioconductor packages (www.bioconductor.org).

Gene Set Enrichment Analysis (GSEA) was conducted using GSEA software [18].

Data Repository

This data has been deposited at the public repository Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) with accession number GSE11903.

Supporting Information

Table S1 DEG genes identified in this study (FDR<0.05, FCH>2).

Table S2 Genes consistently identified by the four studies.

Acknowledgments

We thank Dr. J. Mee for sharing his genomic data with our group. We thank Irna Cardinale and RU Genomic facility for the hybridization of Affymetrix chips. We thank Dr. Marcelo O. Magnasco and Dr. Leanne M. Huang-Johnson for valuable comments and suggestions.

Author Contributions

Conceived and designed the experiments: MSF JGK. Performed the experiments: MAL LCZ. Interpreted data analysis: JGK. Interpreted results: MAL LCZ.

References

1. Irizarry RA, Warren D, Spencer F, Kim IF, Biswal S, et al. (2005) Multiple-laboratory comparison of microarray platforms. Nat Methods 2: 345–350.
2. Larkin JE, Frank BC, Gavras H, Sultana R, Quackenbush J (2005) Independence and reproducibility across microarray platforms. Nat Methods 2: 337–343.
3. Shi LM, Perkins RG, Fang H, Tong WD (2008) Reproducible and reliable microarray results through quality control: good laboratory proficiency and appropriate data analysis practices are essential. Current Opinion in Biotechnology 19: 10–18.
4. Suarez-Farinas M, Magnasco MO (2007) Comparing microarray studies. Methods Mol Biol 377: 139–152.
5. Suarez-Farinas M, Noggle S, Heke M, Hemmati-Brivanlou A, Magnasco MO (2005) Comparing independent microarray studies: the case of human embryonic stem cells. BMC Genomics 6: 99.
6. Shi J, Walker MG (2007) Gene set enrichment analysis (GSEA) for interpreting gene expression profiles. Current Bioinformatics 2: 133–137.
7. Abatangelo L, Magletta R, Distaso A, D’Addabbo A, Cireanza TM, et al. (2009) Comparative study of gene set enrichment methods. BMC Bioinformatics 10: 275.
8. Nam D, Kim SY (2008) Gene-set approach for expression pattern analysis (vol 9, pg 189, 2008). Briefings in Bioinformatics 9: 430–450.
9. Mosoth V, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, et al. (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 34: 267–273.
10. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 102: 15545-15550.
11. Bowcock AM, Shannon W, Du F, Duncan J, Cao K, et al. (2001) Insights into psoriasis and other inflammatory diseases from large-scale gene expression studies. Hum Mol Genet 10: 1793–1803.
12. Gudjonsson JE, Ding J, Li X, Nair RP, Tejayvi T, et al. (2009) Global gene expression analysis reveals evidence for decreased lipid biosynthesis and increased innate immunity in uninvolved psoriatic skin. J Invest Dermatol 129: 2793–2804.
13. Oestreicher JL, Walters IB, Kikuchi T, Gilleaudeau P, Surette J, et al. (2001) Molecular classification of psoriasis disease-associated genes through pharmacogenomic expression profiling. Pharmacogenomics J 1: 272–287.
14. Yao Y, Richman L, Morehouse C, de los Reyes M, Higgs BW, et al. (2008) Type I interferon: potential therapeutic target for psoriasis? PLoS One 3: e2737.
15. Zhou X, Krueger JG, Kao MC, Lee E, Du F, et al. (2003) Novel mechanisms of T-cell and dendritic cell activation revealed by profiling of psoriasis on the 63,100-element oligonucleotide array. Physiol Genomics 13: 69–70.
16. Zaba LC, Cardinale I, Gilleaudeau P, Sullivan-Whalen M, Farinas MS, et al. (2007) Amelioration of epidermal hyperplasia by TNF inhibition is associated with reduced Th17 responses. J Exp Med 204: 3183–3194.
17. Zaba LC, Suarez-Farinas M, Fuentes-Duculan J, Nograles KE, Guttman-Yassky E, et al. (2009) Effective treatment of psoriasis with etanercept is linked to suppression of IL-17 signaling, not immediate response TNF genes. J Allergy Clin Immunol 124: 1022-1030 e1021–1395.
18. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP (2007) GSEA-P: a desktop application for Gene Set Enrichment Analysis. Bioinformatics 23: 3251–3253.
19. Segal E, Shapiro M, Regev A, Pe’er D, Botstein D, et al. (2003) Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. Nat Genet 34: 166–176.
20. Nograles KE, Zaba LC, Guttman-Yassky E, Fuentes-Duculan J, Suarez-Farinas M, et al. (2008) Th17 cytokines interleukin (IL)-17 and IL-22 modulate distinct inflammatory and keratinocyte-response pathways. Br J Dermatol 159: 1086–1091.
21. Mee JB, Johnson CM, Morar N, Burden D, Groves RW (2007) The psoriatic transcriptome closely resembles that induced by interleukin-1 in cultured keratinocytes: dominance of innate immune responses in psoriasis. Am J Pathol 171: 32–42.
22. Shen F, Gaffen SL (2008) Structure-function relationships in the IL-17 receptor: implications for signal transduction and therapy. Cytokine 41: 92–104.
24. Gudjonsson JE, Aphale A, Grachtchouk M, Ding J, Nair RP, et al. (2009) Lack of evidence for activation of the hedgehog pathway in psoriasis. J Invest Dermatol 129: 635–640.

25. Homey B, Alemius M, Müller A, Soto H, Bowman EP, et al. (2002) CCL27-CCR10 interactions regulate T cell-mediated skin inflammation. Nature Medicine 8: 157–165.

26. Boury-Jamot M, Sougrat R, Tailhardat M, Le Varlet B, Bonte F, et al. (2006) Expression and function of aquaporins in human skin: Is aquaporin-3 just a glycerol transporter? Biochim Biophys Acta 1758: 1034–1042.

27. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, et al. (2006) The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science 313: 1929–1935.

28. Haider AS, Potem SB, Kaporis H, Cardinale I, Fri J, et al. (2006) Genomic analysis defines a cancer-specific gene expression signature for human squamous cell carcinoma and distinguishes malignant hyperproliferation from benign hyperplasia. J Invest Dermatol 126: 809–811.

29. Kaporis HG, Guttman-Yassky E, Lows MA, Haider AS, Fuentes-Duculan J, et al. (2007) Human basal cell carcinoma is associated with Foxp3+ T cells in a Th2 dominant microenvironment. J Invest Dermatol 127: 2391–2398.

30. Irizarry RA, Wu Z, Jaffee HA (2006) Comparison of Affymetrix GeneChip expression measures. Bioinformatics 22: 789–794.

31. Shi L, Jones WD, Jensen RV, Harris SC, Perkins RG, et al. (2008) The balance of reproducibility, sensitivity, and specificity of lists of differentially expressed genes in microarray studies. BMC Bioinformatics 9 Suppl 9: S10.

32. Tan PK, Downey TJ, Spitznagel EL Jr., Xu P, Fu D, et al. (2003) Evaluation of gene expression measurements from commercial microarray platforms. Nucleic Acids Res 31: 5676–5684.

33. Zhang M, Yao C, Guo Z, Zou J, Zhang L, et al. (2008) Apparently low reproducibility of true differential expression discoveries in microarray studies. Bioinformatics 24: 2057–2063.

34. Jiang Z, Gentleman R (2007) Extensions to gene set enrichment. Bioinformatics 23: 306–313.

35. Efron B, Tibshirani R (2007) On testing the significance of sets of genes. Annals of Applied Statistics 1: 107–129.

36. Dinu I, Potter JD, Mueller T, Liu Q, Adewale AJ, et al. (2009) Gene-set analysis and reduction. Brief Bioinform 10: 24–34.

37. Ackermann M, Strimmer K (2009) A general modular framework for gene set enrichment analysis. BMC Bioinformatics 10: 47.