

## ORIGINAL ARTICLE

# Integrated computational approach to the analysis of RNA-seq data reveals new transcriptional regulators of psoriasis

Alena Zolotarenko<sup>1</sup>, Evgeny Chekalin<sup>1</sup>, Alexandre Mesentsev<sup>1</sup>, Ludmila Kiseleva<sup>2</sup>, Elena Griбанова<sup>2</sup>, Rohini Mehta<sup>3</sup>, Ancha Baranova<sup>3,4,5,6</sup>, Tatiana V Tatarinova<sup>6,7,8</sup>, Eleonora S Piruzian<sup>1</sup> and Sergey Bruskin<sup>1,5</sup>

Psoriasis is a common inflammatory skin disease with complex etiology and chronic progression. To provide novel insights into the regulatory molecular mechanisms of the disease, we performed RNA sequencing analysis of 14 pairs of skin samples collected from patients with psoriasis. Subsequent pathway analysis and extraction of the transcriptional regulators governing psoriasis-associated pathways was executed using a combination of the MetaCore Interactome enrichment tool and the cisExpress algorithm, followed by comparison to a set of previously described psoriasis response elements. A comparative approach allowed us to identify 42 core transcriptional regulators of the disease associated with inflammation (NFκB, IRF9, JUN, FOS, SRF), the activity of T cells in psoriatic lesions (STAT6, FOXP3, NFATC2, GATA3, TCF7, RUNX1), the hyperproliferation and migration of keratinocytes (JUN, FOS, NFIB, TFAP2A, TFAP2C) and lipid metabolism (TFAP2, RARA, VDR). In addition to the core regulators, we identified 38 transcription factors previously not associated with the disease that can clarify the pathogenesis of psoriasis. To illustrate these findings, we analyzed the regulatory role of one of the identified transcription factors (TFs), FOXA1. Using ChIP-seq and RNA-seq data, we concluded that the atypical expression of the FOXA1 TF is an important player in the disease as it inhibits the maturation of naive T cells into the (CD4+FOXA1+CD47+CD69+PD-L1(hi)FOXP3-) regulatory T cell subpopulation, therefore contributing to the development of psoriatic skin lesions.

*Experimental & Molecular Medicine* (2016) 48, e268; doi:10.1038/emm.2016.97; published online 4 November 2016

## INTRODUCTION

Psoriasis is a common chronic immune-mediated inflammatory condition characterized by complex alterations of cell signaling leading to the progression of the disease. The observed synergy between the aberrant activation of immune cells and the abnormal proliferation and differentiation of keratinocytes leads to the development of typical psoriatic symptoms—red scaly thickened plaques on the skin surface. Another feature of psoriasis is a ‘cytokine storm’ that begins locally within the skin and then spreads throughout the body in form of systemic inflammation that contributes to the development of comorbidities, such as heart disease, stroke, diabetes and psoriatic arthritis.

To identify the key signaling cascades and gene expression alterations causing disease development and progression, we

performed RNA-sequencing (RNA-seq) analysis of the skin transcriptome in 14 patients with psoriasis. Compared with other methods of gene expression analysis, RNA-seq provides a more precise measurement of transcription levels, a wider dynamic range of detection and more reproducible results. It was noted in Quigley<sup>1</sup> that while for the most abundant transcripts both microarray and RNA-seq produce similar results, RNA-seq is capable of identifying a large number of transcripts expressed at low levels that could not be confidently called as differentially expressed when using microarrays to analyze the same number of samples.

In this study, we present the results of RNA-seq analysis that allowed us to identify important signaling cascades enriched with differentially expressed genes (DEGs) and highlight potential transcriptional regulators contributing to

<sup>1</sup>Laboratory of Functional Genomics, Vavilov Institute of General Genetics, Moscow, Russian Federation; <sup>2</sup>Bryansk Regional STD and Dermatology Center, Bryansk Region, Russian Federation; <sup>3</sup>School of Systems Biology, George Mason University, Fairfax, Virginia, USA; <sup>4</sup>Russian Centre for Medical Genetics RAMN, Moscow, Russian Federation; <sup>5</sup>Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region, Russian Federation; <sup>6</sup>Atlas Oncology Diagnostics, Moscow, Russian Federation; <sup>7</sup>Center for Personalized Medicine, Children’s Hospital Los Angeles and Spatial Sciences Institute, University of Southern California, Los Angeles, California, USA and <sup>8</sup>AA Kharkevich Institute for Information Transmission Problems RAS, Moscow, Russian Federation  
Correspondence: Dr S Bruskin, Laboratory of Functional Genomics, Vavilov Institute of General Genetics, Gubkina Street 3, Moscow 119991, Russian Federation.

E-mail: sergey.bruskin@gmail.com

Received 10 March 2016; revised 6 May 2016; accepted 24 May 2016

the development of the disease. To identify transcriptional regulators of psoriatic pathology, we utilized two knowledge-based tools, MetaCore<sup>2</sup> and cisExpress.<sup>3,4</sup> Modulation of the identified signaling pathways may be a promising approach for the development of novel management strategies of psoriasis and other diseases commonly associated with this condition.<sup>5-7</sup>

## MATERIALS AND METHODS

### Ethics statement

All steps of data collection and analysis were conducted according to the Declaration of Helsinki principles. Informed written consent was obtained from human subjects under protocols approved by the Ethics Committee of the Vavilov Institute of General Genetics, Russian Academy of Sciences.

### Patients and samples

The patients in this study were unrelated Caucasian individuals with the plaque form of psoriasis from the Bryansk regional STD and Dermatology Center. The metadata for the patients are presented in Supplementary Table S1. Two 4-mm punch biopsy specimens were taken from the skin of the patients, one from the lesional area of the skin (LS sample) and another from the non-lesional area of skin 3–4 cm apart from the lesion, in an area that did not have any visual signs of psoriasis (NL sample). Patients did not obtain any systemic or PUVA/UV treatment for 1 month before the biopsies. All biopsy samples were immediately transferred to the liquid nitrogen until RNA extraction.

### RNA sequencing

A TissueLyser LT homogenizer (Qiagen, Valencia, CA, USA) was used to homogenize the biopsy specimens. Total RNA was extracted with the ExtractRNA reagent (Evrogen, Moscow, Russia) according to the manufacturer's protocol. Isolated RNA was dissolved in RNase-free water, rRNA was depleted using a RiboMinus Eukaryote Kit for RNA-seq (Life Technologies, Thermo Fisher Scientific, Waltham, MA, USA) and the samples were stored at  $-80^{\circ}\text{C}$ . The quality of total RNA was evaluated with an RNA 6000 Pico Chip Kit and an Agilent 2100 Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA, USA) and with a Quant-iT RNA Assay Kit and a Qubit fluorometer (Life Technologies). The average RNA integrity number of the samples was  $\geq 7$ . Library preparation and sequencing were performed using a SOLiD 4 System platform and sequencing chemistry according to the manufacturer's instructions (Life Technologies).

### Processing and mapping of RNA-seq reads and differential expression analysis

Raw pair-end reads (50+25 bp) were obtained from the SOLiD4 System (Applied Biosystems) in color space format (\*.csfasta) and were filtered for quality. The adaptor sequences were trimmed, and the reads were aligned to the UCSC human reference genome (hg19) using the Applied Biosystems Bioscope software (Applied Biosystems Bioscope, Thermo Fisher Scientific, Waltham, MA, USA) to obtain reads in the BAM format. Mapping to multiple locations was permitted. The aligned read BAM files were assembled into transcripts, their abundance was estimated and tests for differential expression were processed using the Bioconductor DESeq package (Simon Anders, EMBL, Heidelberg, Germany).<sup>8</sup> False discovery rate

(FDR) correction for multiple testing was performed according to Benjamini *et al.*<sup>9-11</sup>

### Pathway analysis and identification of transcriptional regulators

Using a fold change cutoff (FC)  $> 1.5$  and a FDR  $< 0.05$  and including only those genes that had reads at all samples, we identified 1564 DEGs: 938 of them were upregulated and 626 were downregulated. This subset of DEGs was used for gene ontology (GO) analysis using the DAVID tool (Database for Annotation, Visualization and Integrated Discovery ver. 6.7)<sup>12</sup> and pathway analysis as well as interactome analysis using the MetaCore database from Thomson Reuters (ver. 6.11, build 41105, GeneGo, Thomson Reuters, New York, NY, USA).<sup>2</sup> The MetaCore Pathway analysis tool was used to perform gene network enrichment analysis; the MetaCore Interactome tool was used for the identification of transcriptional regulators of DEG-enriched pathways.<sup>2</sup> The transcriptional regulators of DEGs were ranked by Z-score (representing the level of connectivity) and calculated as follows:

$$Z - \text{score} = \frac{r - n \frac{R}{N}}{\sqrt{n \left(\frac{R}{N}\right) \left(1 - \frac{R}{N}\right) \left(1 - \frac{n-1}{N-1}\right)}} = \frac{r - \mu}{\sigma}, \text{ where}$$

$N$  is the total number of nodes (genes) after filtration;  $R$  is equal to either the number of nodes in the input list or the nodes associated with the experimental data;  $n$  is the number of the nodes in the network;  $r$  is the number of the network's nodes associated with the experimental data or included in the input list;  $\mu$  and  $\sigma$  are the mean and dispersion of the corresponding hypergeometric distribution, respectively.

The cisExpress algorithm<sup>3,4</sup> was used for the identification of promoter motifs and the discovery of cis-elements in promoter sequences that are statistically associated with the expression patterns of DEG. Promoter sequences were obtained from the EPDnew database,<sup>13</sup> which is a collection of experimentally validated promoters in the human, mouse, fruit fly and zebrafish genomes. Evidence comes from transcription start site mapping from high-throughput experimental techniques, such as CAGE<sup>14</sup> and Oligocapping.<sup>15</sup> The positions of the promoters were validated using the NPEST algorithm.<sup>16</sup> We identified 16 542 *Homo sapiens* promoters with corresponding RNA-seq gene expression measurements in lesional and non-lesional skin.

The relative expression (R.E.) values for every gene were calculated from average gene expression data for lesional ( $n=14$ ) and non-lesional ( $n=14$ ) skin according to the formula:

$$R.E. = \ln \left( \frac{\langle \text{non-lesional expression} \rangle}{\langle \text{lesional expression} \rangle} \right).$$

The Z-score for cisExpress data were calculated as follows:

$$Z - \text{score}(w, k) = \frac{e_{\text{with}}(w, k) - e_{\text{without}}(w, k)}{\sqrt{\frac{Stdev_{\text{with}}^2(w, k)}{n_{\text{with}}(w, k)} + \frac{Stdev_{\text{without}}^2(w, k)}{n_{\text{without}}(w, k)}}}, \text{ where}$$

$e_{\text{with}}(w, k)$  and  $e_{\text{without}}(w, k)$  are average gene expression values;  $Stdev_{\text{with}}(w, k)$  and  $Stdev_{\text{without}}(w, k)$  are s.d. of gene expression values; and  $n_{\text{with}}(w, k)$  and  $n_{\text{without}}(w, k)$  are the number of sequences of genes containing and not containing the motif  $w$  in the  $k$ th window.

### FOXA1 target gene identification in the DEG list

The FOXA1 target list was obtained by merging ChIP-seq data from GSE39241<sup>17</sup> and GSM1099031<sup>18</sup> using the edgeR<sup>19</sup> and DESeq<sup>8</sup>

packages with default parameters and then cross-referencing the results with the DEG list.

## RESULTS

### Analysis of psoriasis-associated DEGs

In this study, we identified 1564 genes that were differentially expressed in psoriatic lesions (psoriatic DEGs): 938 of them were upregulated and 626 were downregulated. Analysis of the top 20 upregulated DEGs (Supplementary Table S2) highlighted the importance of immune defense mechanisms, inflammatory response, taxis and chemotaxis of immune cells and alterations of epidermal differentiation in the pathogenesis of psoriasis. Our analysis suggests that the genes with the largest magnitude of expression changes are the ‘response’ genes that contribute to the pathophysiological manifestations of psoriasis rather than the initiation of the disease. A majority of the top 20 downregulated genes (Supplementary Table S3) were linked to lipid biosynthesis and lipid metabolism. Interestingly, among the top 20 downregulated genes, we detected a number of poorly characterized expression units, including possible pseudogenes, and non-coding RNAs. It remains an open question whether non-coding RNA enrichment is of any functional significance or just an indication of technology bias.

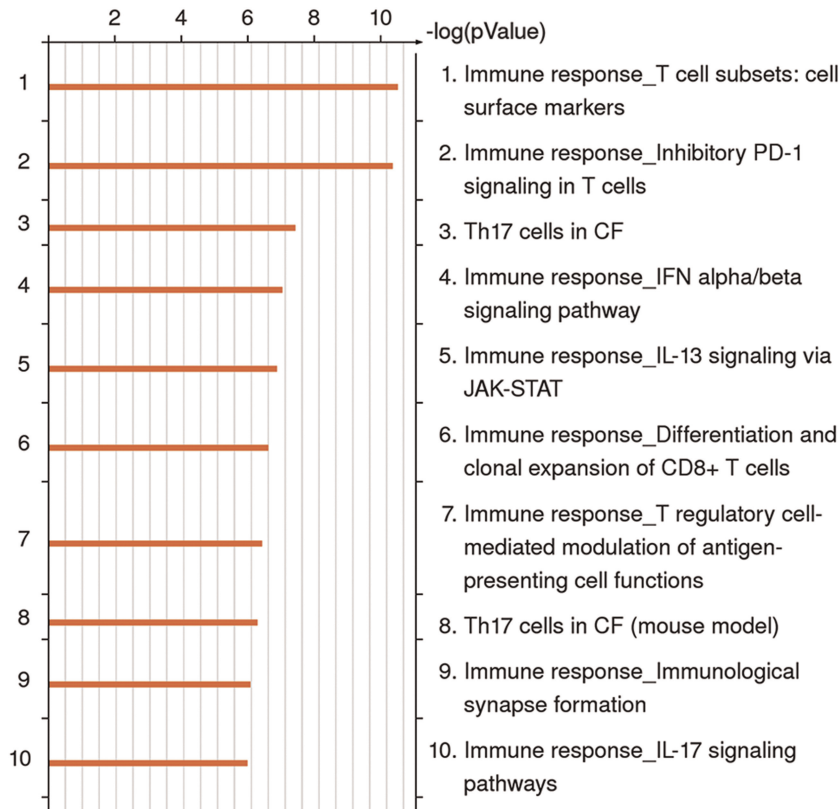
To identify the molecular basis of psoriatic pathology, GO analysis and MetaCore-guided pathway analysis were performed.<sup>2</sup> The results of the GO analysis (Supplementary

Table S4) generally supported the findings obtained with MetaCore pathway enrichment (Figure 1) and the analysis of top DEGs described above. A detailed discussion of a number of psoriasis-associated pathways is presented in the ‘Discussion’ and ‘Supplementary Materials’ sections.

### Transcriptional regulation

The skin serves as a first line of defense against pathogen invasion. The stimulation of different pathogen-sensing receptors (such as pathogen-recognition receptors) leads to the activation of antimicrobial defense (i.e. defensins and other gene clusters, Supplementary Figures S1 and S2, Supplementary Table S5) that is orchestrated by a number of key transcription factors (TFs), including nuclear factor kappa B (NFκB), activator protein 1, cAMP response element-binding protein, interferon-regulatory factors (IRF) and others. Although the putative antigen leading to the activation of pro-inflammatory signaling in psoriasis has not yet been identified, it is widely accepted that the signaling cascades activated in the course of psoriatic inflammation are mainly the same as those stimulated during pathogen invasion. The activation of inflammatory and antiapoptotic proteins ultimately alerts the immune system of the invasion and induces the recruitment of leukocytes to the site of infection.<sup>20</sup>

To ascertain the key regulatory ‘hub points’ of the psoriatic networks, two independent approaches were used. The first of



**Figure 1** Top 10 DEG-enriched signaling pathways. Sorted by statistical significance of the findings. The results were obtained using the MetaCore pathway analysis tool (GeneGO/Thomson Reuters).

them utilizes the MetaCore Interactome enrichment tool<sup>2</sup> that evaluates the levels of connectivity between the nodes (that can be either proteins or genes), identifies overconnected nodes and, according to the node function, suggests possible transcriptional regulators that drive the observed pattern of gene expression in the entire data set. The second approach relies on the cisExpress algorithm,<sup>3,4</sup> which allows one to perform *de novo* discovery of the motif within the putative promoter regions of DEGs by means of comparison of these regions with the content of the HOCOMOCO<sup>21</sup> v9, JASPAR,<sup>22</sup> HumanTF 1.0<sup>23</sup> and footprintDB<sup>24</sup> databases of known TF-binding sites (TFBSs). Next the identified lists of transcriptional regulators were compared with the data compiled by Swindell *et al.*<sup>25</sup> using meta-analysis of transcriptomes of 237 psoriatic patients, a dictionary of 2935 putative TFBSs and the sites for unconventional DNA-binding proteins. Swindell *et al.*<sup>25</sup> identified psoriasis-response elements that were over-represented upstream of psoriasis DEGs in putative promoters that were defined as sequences starting at 5 kb upstream and ending at 500 bp downstream from the major transcription start site.

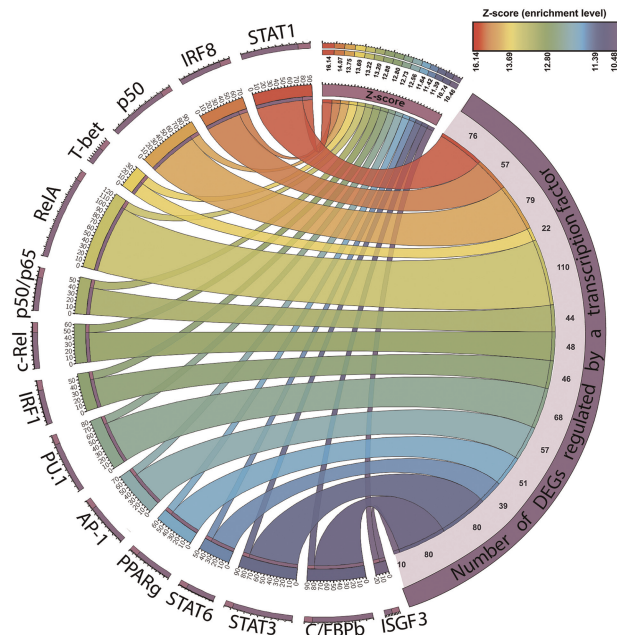
*MetaCore-guided identification of transcriptional regulators.* The MetaCore Interactome tool determines the density of interactions between each protein from a data set of interest

and all other proteins, evaluates statistically significant interactions within the set and analyzes the functions of the selected interacting proteins.<sup>2</sup> Because proteins usually work in groups (such as protein complexes and pathways) that are defined by protein interactions, it is assumed that the relative connectivity of each hub reflects its relevance, or importance, and it may be used for the identification of transcriptional regulators of DEG-enriched signaling cascades. Even if the expression levels of the mRNA encoding the TF itself are not altered, for example, when the TF in question is predominantly regulated posttranscriptionally, the number of targets it interacts with depends on the state of its activation or suppression. Hence, the enrichment or the depletion of an interacting protein network indicates the activation or suppression of the TF that orchestrates the network.

We identified possible transcriptional regulators of DEGs (the TFs significantly associated with DEG list; Supplementary Table S6) and computed their *Z*-score—the degree of connectivity of the TF to the list of DEGs. The *Z*-score signifies a number of connections obtained between a TF and a list of DEGs in excess of the number of connections expected with a random set of genes of the same size. The identified ‘Top’ transcriptional regulators of the DEGs (Figure 2) are ranked according to their *Z*-scores corresponding to the influence of the TFs on the pathological processes associated with the disease. Their regulation leads to the development of the main distinctive features of psoriasis. In the ‘Top’ TF list, there are both cell type-specific TFs (for example, PU.1, which is a master regulator of myeloid cells<sup>26</sup>) and ubiquitously expressed TFs associated with inflammatory pathways (for example, NFκB and IRF<sup>27</sup>), which reflects alterations in cell populations in a plaque compared with the unaffected skin. In addition to the TFs commonly associated with psoriasis, we identified TFs that have not been previously associated with the disease (Supplementary Table S7).

Further analysis is needed to determine the roles of the identified DEGs in disease progression. To illustrate our hypothesis of their involvement in psoriasis, we analyzed one of the TFs, FOXA1, and its target genes, looking for possible associations with pathogenesis. Using the CHIP-seq data available from experiments GSE39241<sup>17</sup> and GSM1099031,<sup>18</sup> we have compiled a list of gene targets for this TF and compared it with the list of DEGs identified by RNA-seq (Supplementary Table S12). The intersection between the two lists was used in the GO enrichment analysis to reveal processes in psoriatic lesions regulated by this TF (Supplementary Table S13). FOXA1 was found to be associated with the immune alterations indicative for the disease as well as with the abnormalities of lipid metabolism.

*De novo analysis of transcriptional regulation of DEGs using the cisExpress algorithm.* Another approach that we utilized for the identification of the transcriptional regulators of DEGs was based on the cisExpress algorithm.<sup>3,4</sup> cisExpress finds putative regulatory elements using a combination of sequence and expression information. We used 16 542 validated *H. sapiens*



**Figure 2** Top 15 transcriptional regulators of DEG genes (identified with MetaCore Interactome). Transcription factors are ranked according to their *Z*-score (the level of connectivity of the TF to the DEG list). Larger *Z*-scores represent higher levels of connectivity between the transcription factor and the DEG list. The colors from red to violet indicate *Z*-score values from higher (16.14) to lower (10.48). The width of the ribbon for the transcription factor corresponds to the number of target genes for the transcription factor in the DEG list (for example, for the *STAT1* gene, the *Z*-score is 16.14, and the number of regulated DEGs is 76). The numerical data for this figure are in Supplementary Table S6.

promoters that have corresponding RNA-seq gene expression measurements in lesional and non-lesional skin.

The length of a promoter region varies from gene to gene, and the identification of a 'promoter window' containing the most important regulatory sequences for each gene is a separate challenge. Hence, for our analysis, we used the 'core promoter-5' untranslated region' of (+500, -500) base pairs around the transcription start site of each gene. The analyzed set of the promoters was examined for the presence of motifs (putative TFBSs), and the corresponding gene expression values were compared for genes whose promoters did and did not contain the motifs using a *t*-test. We compiled a ranked list of > 100 position-specific motifs in the promoter regions of DEGs (Supplementary Table S8).

The top 10 motifs in the ranked list have the highest influence on gene expression (Table 1). For every motif, a *Z*-score was calculated. Positive *Z*-score values suggest that the presence of the motif is associated with elevated levels of gene expression in non-lesional skin, while negative values show that the presence of the motif is associated with elevated levels of gene expression in lesional skin. The absolute value of the *Z*-score is used to calculate the confidence level of the influence of the motif on gene expression. The complete list of motifs is given in Supplementary Table S9.

The motifs were examined for similarity with known TFBSs. HOMOCO v9,<sup>21</sup> JASPAR 2014,<sup>22</sup> HumanTF 1.0<sup>23</sup> and footprintDB<sup>24</sup> databases were used for this analysis (see Supplementary Table S9). For example, one of the identified motifs (AAGATG) is related to the ETS1 TF, a negative regulator of Th17 cells,<sup>28</sup> which are one of the most important T-cell populations in psoriasis, and to GATA-3, a transcriptional regulator leading to T helper type 2 (Th2) polarization in the Th1/Th2-cell-type switch,<sup>29</sup> which is downregulated in psoriatic skin. The CCGGAA motif is associated with the TF ELK4, which is highly expressed in lesional psoriatic skin.<sup>30</sup>

*Identification of the key transcriptional regulators of the psoriatic transcriptome.* To find the key transcriptional regulators of the DEGs, we compared lists identified by the two computational approaches (MetaCore and cisExpress) and the results of Swindell *et al.*<sup>25</sup> Comparison of the three groups of transcriptional regulators (327 cisExpress-identified, 200 MetaCore-identified and 439 identified by Swindell *et al.*) (Figure 3) identified 42 common TFs representing the 'core' TF regulators of the psoriatic transcriptome (Table 2).

The majority of elements of the 'core TF' list are TFs associated with inflammation (NFκB, IRF9, JUN, FOS, SRF), the activity of T-cells in the psoriatic lesions (STAT6, FOXP3, NFATC2, GATA3, TCF7, RUNX1 etc.), hyperproliferation and migration of keratinocytes (JUN, FOS, NFIB, TFAP2A, TFAP2C) and lipid metabolism (TFAP2, RARA, VDR). There were several FOX ('fork-head' box) family proteins in the list containing the evolutionary conserved 'fork-head' or 'winged-helix' DNA-binding domain. These proteins could work as active regulators of cell proliferation and metabolism and also serve as pioneer factors that

de-condense chromatin, therefore facilitating the binding of other sequence-specific TFs to target enhancers, repressors and promoters, wiring global gene networks essential for cell fate decisions.<sup>31</sup> We also found 294 DEG-associated TFs not identified by Swindell *et al.*<sup>25</sup> Sixty-six of them were identified by cisExpress as well as by MetaCore (Figure 3 and Supplementary Table S10). Among the negative regulators, there is the CAGAC motif, frequently found in Smad-responsive promoter regions. Smad TFs have a key role in cytokine signaling pathways. The Smad complex includes DNA-binding cofactors influencing the recruitment of transcriptional coactivators or corepressors.<sup>32</sup> Another TFBS with a negative association with gene expression in psoriatic lesions is the CACCC-box, a binding site for the Krüppel-like factors that can act as a transcriptional repressor.<sup>33–37</sup> Recently, Krüppel-like factors were reported to regulate epidermal proliferation and differentiation in humans and control the physiological reaction to counteract the abnormal differentiation and proliferation of keratinocytes.<sup>38</sup>

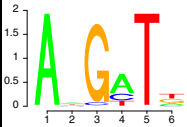
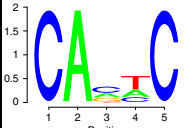
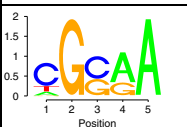
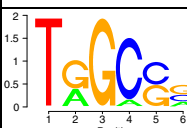
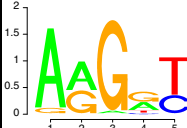
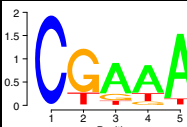
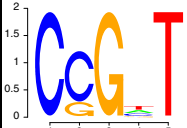

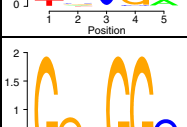
### Pathway analysis

To investigate the relationships between the DEGs, we performed gene network enrichment analysis using the MetaCore software. The top 10 signaling networks enriched with DEGs were mainly associated with different alterations in immune signaling present in the psoriatic lesions (Figure 1), for example, the map 'Immune response\_IL-17 signaling pathway' (Figure 4). This agrees with the hypothesis that the main feature of psoriasis is the cytokine storm and altered balance of cytokines, chemokines and growth factors regulating various immune and inflammatory responses (see Supplementary Materials for a detailed discussion). The main psoriasis-associated transcriptional regulators of this map are NFκB and C/EBPβ, as defined by the overexpression of their target genes shown on the map (for example, *CCL2*, *CCL7*, *CCL20*, *HBD2*, *IL8*, *iNOS2* and so on) and the consequent activation of downstream processes important for disease pathogenesis (chemotaxis of neutrophils, dendritic and T cells, alterations in cell adhesion and antibacterial activity). All of the TFs on the map were identified in the Top15 list in MetaCore analysis (Figure 2) as well as in the cisExpress analysis (Supplementary Table S8).

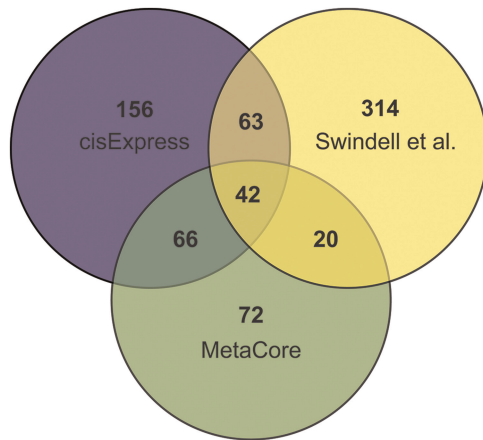
### DISCUSSION

In psoriasis, the observed changes in the gene expression levels and the DEG enrichment of certain GO processes may be explained by two different disease-associated phenomena: changes in the transcription and degradation rates of mRNA and alterations in the composition of cells within the lesion, which is usually characterized by epidermal thickening, the accumulation of immune cells and thinning of the subcutaneous fat layer. Therefore, whether the differential expression observed between lesional and non-lesional psoriatic samples truly reflects alterations in intracellular signaling remains unclear.<sup>37</sup> The top 10 signaling pathways (Figure 1) highlighted the importance of the activation and chemotaxis of

**Table 1 Top 10 motifs with the highest confidence of influence on gene expression identified with cisExpress**

From...to, bp	Motif	Z-score	Source	E-value	Associated proteins	
10...30	AAGATG		7.08	1	1.3e-05	ETS1, p54
				2	1.4e-05	AP-1, p39, AP1
-20...0	CCGGAA	5.87	3	3.1e-08	ELK4	
-10...10	CAC[CT]C		-5.71	4	2.3e-06	ZIC1, ZIC2, ZIC3
				4	6.1e-06	GLI3, KLF1
				2	1.3e-05	SREBF1
				4	3.8e-04	NKX25
-60...-40	CGGAA		5.64	1	2.7e-08	NFATC2, ETS2, ELK4
				4	5.8e-06	NRF-2/GABP1
				3	5.9e-06	ELK1
10...30	TGGCGG		5.64	3	8.6e-08	E2F8, NF-E1
				4	1.7e-06	TFDP1
0...20	AAGAT		5.56	3	1.9e-05	TCF7L1
				4	0.001	GATA2, GATA6
				2	0.001	JUN
-50...-30	CGGAA		5.44	1	2.7e-08	NFATC2, ETS2, ELK4
				4	5.8e-06	NRF-2/GABP1
				3	5.9e-06	ELK1
300...320	CCGGT		5.43	4	6.0e-06	ELK4
				3	6.0e-06	GRHL1
-10...10	GCCAT		5.37	4	6.9e-05	RFX3
				3	8.1e-05	E2F2
0...20	GATGGC		5.31	4	8.3e-07	ZBTB4
				4	4.9e-05	HXA1, HXB1
				4	1.5e-04	TAL1

'From...to' is the position of the 'window' where the motif was discovered relative to the gene transcription start. Source: 1, footprintDB;<sup>20</sup> 2, JASPAR;<sup>18</sup> 3, HumanTF 1.0;<sup>19</sup> 4, HOCOMOCO.<sup>17</sup> Positive Z-score values suggest that the presence of the motif promotes the expression of a gene in non-lesional skin, while negative Z-score values suggest that the motif acts in lesional skin.



**Figure 3** Venn diagram showing overlap between lists of transcriptional regulators of DEG. green, identified by the MetaCore Interactome software; Violet, TFs identified by the cisExpress tool; yellow, identified in Swindell *et al.*<sup>25</sup>

immune cells mediated by the locally enhanced production of pro-inflammatory cytokines and chemokines and increasing inflammation. The importance of different populations of T cells in the pathogenesis of the disease is illustrated by the DEG-enriched map ‘Immune response: T-cell subsets secreted signals’ (Supplementary Figure S3). It shows the shift of T-cell populations toward interleukin-17-producing types, indicating the activation and enhanced migration of psoriasis-specific populations of T cells to the lesional skin.

One of the main regulators of T-cell polarization and population maintenance is the regulatory T-cell (Treg) population (Supplementary Figure S4). Alterations in the presence and activity of Tregs in the skin of psoriasis patients may trigger development of the disease.<sup>39</sup> The canonical transcriptional regulator of this cell population is the TF FOXP3. A recent study described a new population of Treg cells that carry a non-canonical marker, FOXA1, instead of the canonical FOXP3.<sup>40</sup> This population also has a suppressive role in autoimmunity as adoptive transfer of such cells inhibited experimental autoimmune encephalomyelitis in a FOXA1- and PD-L1-dependent manner.<sup>40</sup> Our analysis identified the upregulation of FOXP3 (FC=2.30, FDR=0.001) and downregulation of FOXA1 (FC=0.49, FDR=0.01) in psoriatic lesional skin. We found that most of the markers of Treg cells were overexpressed in the analyzed transcriptomic data (*CD4*, *CD47*, *CD69*, *PD-L1*) except for FOXA1 itself. We hypothesize that the observed reduced expression of FOXA1 represents differences between Treg populations in non-lesional and lesional psoriatic skin. We suggest that the inhibition of maturation of naive T cells into a FOXA1+ Treg subpopulation (*CD4+FOXA1+CD47+CD69+PD-L1(hi)FOXP3-*) contributes to the development of the disease.

The TF FOXA1 is not T-cell specific; it is involved in the regulation of cell differentiation under normal and pathological conditions (in epithelial, prostate and breast cancers; in epithelial cells of the intestines and lungs; and in pancreatic

cells and other tissues<sup>41–47</sup>). During embryonic stem cell development and hepatic differentiation, FOXA1 functions as a ‘pioneer TF’ owing to its ability to engage condensed chromatin, bind nucleosome-assembled FOXA1 regulatory elements and displace repressive linker histones in response to retinoic acid treatment and transforming growth factor- $\beta$  signaling.<sup>48</sup>

In addition to its role in hepatic epithelial differentiation, FOXA1 is also involved in uroepithelial differentiation. Differentiation of the urothelium is a process governed by peroxisome proliferator-activated receptor- $\gamma$  (PPAR- $\gamma$ ) that involves intermediary TFs such as *IRF-1* and FOXA1. siRNA knockdown of either FOXA1 or *IRF-1* abrogates the PPAR- $\gamma$ -induced induction of late/terminal urothelial differentiation-associated genes.<sup>49</sup> PPAR- $\gamma$  is also an important regulator of keratinocyte differentiation as well as the formation of the horny layer of the epidermis, which is rich in lipids and important for the barrier function. Treatment of cultured human keratinocytes with ciglitazone, a PPAR- $\gamma$  activator, increases the mRNA levels of involucrin and transglutaminase-1. Moreover, topical treatment of hairless mice with ciglitazone or troglitazone increases loricrin, involucrin and filaggrin expression. As PPAR $\gamma$  itself is downregulated according to the RNA-seq data (FC=0.39, FDR=0.0039), and it is an upstream regulator of FOXA1 signaling, it could be speculated that aberrant fatty acid metabolism as well as the general metabolic disturbances observed in psoriasis could lead to impaired cell differentiation via the inhibition of PPAR $\gamma$  and the associated inhibition of FOXA1 expression in the skin.

Another possible consequence of reduced FOXA1 expression is disturbed keratinocyte differentiation (more discussion on this topic is in the Supplementary Materials). To evaluate the putative contribution of FOXA1 regulation to the development of the psoriatic process, we have compared lists of FOXA1 targets, identified in ChIP-seq experiments by Hurtado *et al.*,<sup>50</sup> with the DEGs identified by our RNA-seq analysis. We demonstrated that FOXA1 is a transcriptional regulator of the top DEGs that serve as major histopathological contributors (Supplementary Table S2), encoding S100 proteins, serpins and genes for chemoattractant CXCL proteins. Among other important upregulated targets contributing to the disease were *HLA-DPB1* (an HLA class II beta chain paralog expressed in antigen-presenting cells; a risk allele for the disease<sup>51</sup>), keratins 6B and 6C (activation markers of keratinocytes essential for the formation of keratin intermediate filaments that also take part in wound healing) and PPAR $\delta$  (a TF overexpressed in psoriasis that enhances the proliferation of keratinocytes and is induced by *JUNB* in keratinocytes).<sup>52</sup> Among the downregulated targets were genes associated with lipid disturbances observed in psoriatic lesions, including known components of fatty acid metabolism such as the acyl-CoA wax alcohol acyltransferase gene *AWAT2*, the fatty acid elongase gene *ELOVL3*, the fatty acid-binding protein *FABP4* and many others. Ontology analysis of the FOXA1 target DEGs yielded a list of ontologies similar to the whole DEG list discussed above (Supplementary

**Table 2 Core TF regulators of the psoriatic transcriptome**

<i>Ensembl Gene ID</i>	<i>Gene symbol</i>	<i>Transcription factor name</i>
ENSG00000067955	CBFB	Core-binding factor, beta subunit
ENSG00000105516	DBP	D site of albumin promoter (albumin D-box)-binding protein
ENSG00000101412	E2F1	E2F transcription factor 1
ENSG00000164330	EBF1	Early B-cell factor 1
ENSG00000120738	EGR1	Early growth response 1
ENSG00000134954	ETS1	v-ets erythroblastosis virus E26 oncogene homolog 1 (avian)
ENSG00000170345	FOS	v-fos FBJ murine osteosarcoma viral oncogene homolog
ENSG00000125798	FOXA2	Forkhead box A2
ENSG00000111206	FOXM1	Forkhead box M1
ENSG00000150907	FOXO1	Forkhead box O1
ENSG00000118689	FOXO3	Forkhead box O3
ENSG00000128573	FOXP2	Forkhead box P2
ENSG00000049768	FOXP3	Forkhead box P3
ENSG00000107485	GATA3	GATA-binding protein 3
ENSG00000162676	GF11	Growth factor independent 1 transcription repressor
ENSG00000135100	HNF1A	HNF1 homeobox A
ENSG00000101076	HNF4A	Hepatocyte nuclear factor 4, alpha
ENSG00000213928	IRF9	Interferon regulatory factor 9
ENSG00000177606	JUN	Jun oncogene
ENSG00000169926	KLF13	Kruppel-like factor 13
ENSG00000106689	LHX2	LIM homeobox 2
ENSG00000099326	MZF1	Myeloid zinc finger 1
ENSG00000101096	NFATC2	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2
ENSG00000147862	NFIB	Nuclear factor I/B
ENSG00000165030	NFIL3	Nuclear factor, interleukin 3 regulated
ENSG00000109320	NFKB1	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1
ENSG00000143190	POU2F1	POU class 2 homeobox 1
ENSG00000131759	RARA	Retinoic acid receptor, alpha
ENSG00000159216	RUNX1	Runt-related transcription factor 1
ENSG00000186350	RXRA	Retinoid X receptor, alpha
ENSG00000175387	SMAD2	SMAD family member 2
ENSG00000143842	SOX13	SRY (sex determining region Y)-box 13
ENSG00000125398	SOX9	SRY (sex determining region Y)-box 9
ENSG00000112658	SRF	Serum response factor (c-fos serum response element-binding transcription factor)
ENSG00000166888	STAT6	Signal transducer and activator of transcription 6, interleukin-4 induced
ENSG00000162367	TAL1	T-cell acute lymphocytic leukemia 1
ENSG00000071564	TCF3	Transcription factor 3 (E2A immunoglobulin enhancer-binding factors E12/E47)
ENSG00000081059	TCF7	Transcription factor 7 (T-cell specific, HMG-box)
ENSG00000148737	TCF7L2	Transcription factor 7-like 2 (T-cell specific, HMG-box)
ENSG00000137203	TFAP2A	Transcription factor AP-2 alpha (activating enhancer-binding protein 2 alpha)
ENSG00000087510	TFAP2C	Transcription factor AP-2gamma (activating enhancer-binding protein 2gamma)
ENSG00000111424	VDR	Vitamin D (1,25-dihydroxyvitamin D3) receptor

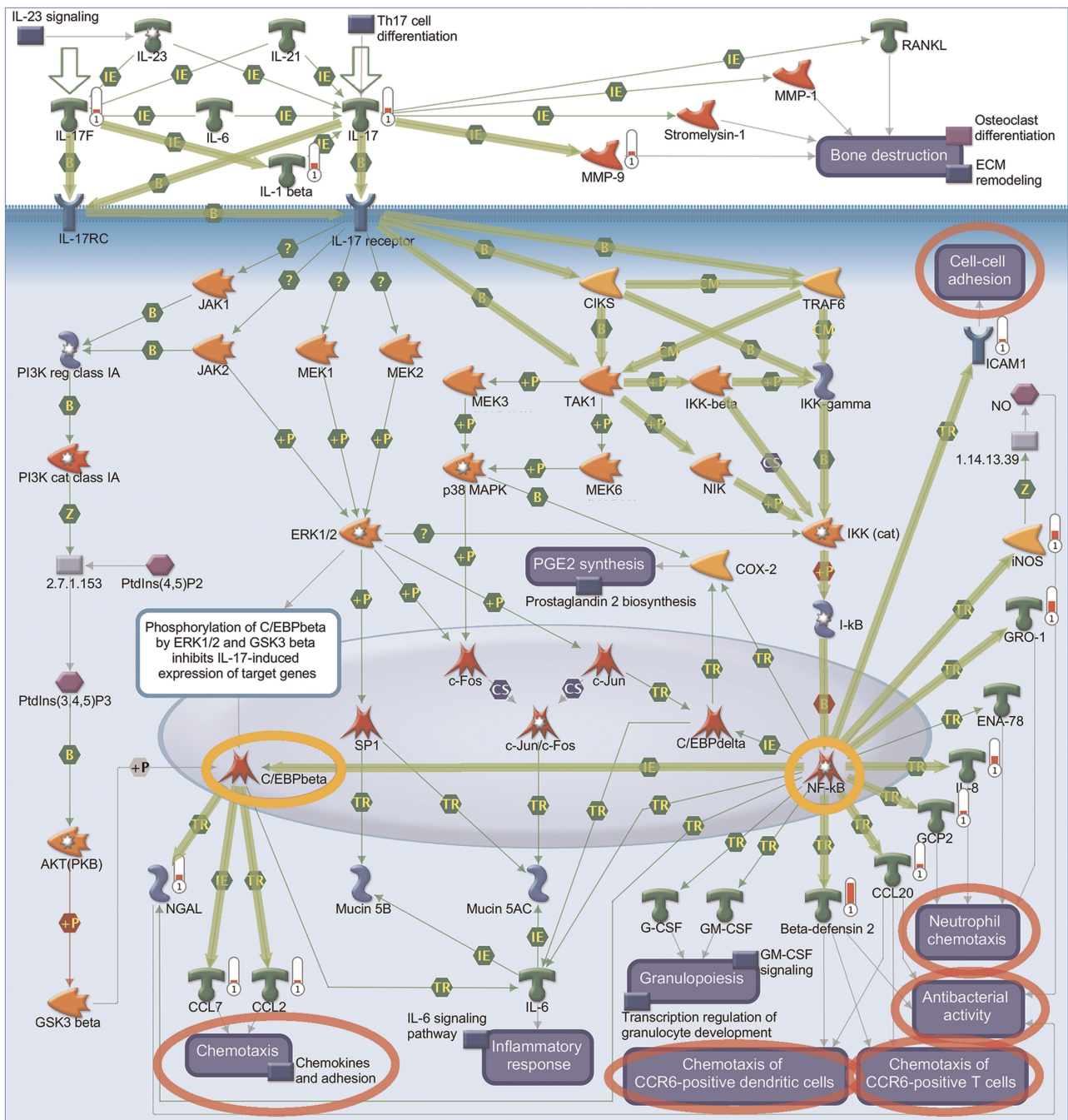
Abbreviation: TF, transcription factor.

Tables S11 and S12). Comparison of the lists of ontologies showed that *FOXA1* is a part of a regulatory complex accounting for the most important psoriatic alterations and signaling cascades important for pathology. Thus this TF is a promising candidate for future investigations of psoriasis.

There are two potential limitations of our study: relatively small samples size (14 pairs) and the ethnic homogeneity of samples (only Caucasians). Earlier studies demonstrated that a sample set of this size is sufficient to identify DEGs.<sup>53,54</sup> Because the RNA sequencing coverage depth corresponds to

a log-normal distribution, a parametric test such as *t*-test could be utilized to identify DEGs.<sup>55</sup> To evaluate the reproducibility of the results, we compared our list of DEGs with a recent large-scale transcriptomic analysis of psoriatic samples (92 samples of lesional skin and 82 samples of non-lesional skin)<sup>37</sup> that identified a larger number of DEGs (6254 DEGs, 4581 upregulated and 1673 downregulated genes, FC>1.5, FDR<0.05). We demonstrated that >60% of the DEGs identified in our study matched the DEGs found by Li *et al.*<sup>37</sup> (Supplementary Figure S6). Given the differences





**Figure 4** Immune response\_IL-17 signaling pathway. Illustration generated with the MetaCore pathway analysis tool (GeneGO/Thomson Reuters) and enriched with DEGs. A description of the map symbols is in Supplementary Figure S5. The red bars near the gene name illustrate the level of gene overexpression. Green arrows indicate signaling cascades enriched with DEGs. Orange circles indicate important transcription regulators of DEGs present in the map. Red circles indicate downstream signaling cascades activated by the pathway in lesional psoriatic skin.

between the sequencing platforms, sample sizes and studied populations, the concordance of >60% is truly remarkable.

To avoid admixture, our samples were obtained from Caucasian patients in the regions where the Caucasian population is predominant. A recent large trans-ethnic genome-wide meta-analysis (15 369 psoriasis cases and 19 517 controls of Caucasian and Chinese ancestries) revealed population-specific

effects or allelic heterogeneity for 11 susceptibility loci and noted that such population-specific effects contribute significantly to the ethnic diversity of psoriasis prevalence and may influence the course of disease.<sup>56</sup> Therefore, our data containing only Caucasian samples allows us to zoom in on immuno-pathogenesis and important psoriasis-associated pathways in Caucasian population. Additional studies should

be conducted to evaluate the applicability of such data to the different ethnic groups.

In conclusion, our analysis highlighted the importance of immune system alterations in the development of psoriasis. We presented a list of identified core transcriptional regulators of the psoriatic transcriptome that can be validated to provide insight into the mechanisms of gene regulation in psoriasis. We also identified novel transcriptional regulators of psoriasis-associated pathways previously not implicated in the pathology. The comparison of our data with public ChIP-seq data allowed us to suggest a hypothesis explaining the role of the TF *FOXA1* in psoriasis.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

The reported study was funded by The Russian Foundation for Basic Research (AZ was supported by 16-34-01356 mol\_a and AM was supported by 14-04-01207A). TVT was supported by grants from The National Institute for General Medical Studies (GM068968), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (HD070996) and the National Science Foundation Division of Evolutionary Biology (1456634).

- 1 Quigley D. RNA-seq permits a closer look at normal skin and psoriasis gene networks. *J Invest Dermatol* 2014; **134**: 1789–1791.
- 2 Bessarabova M, Ishkin A, JeBailey L, Nikolskaya T, Nikolsky Y. Knowledge-based analysis of proteomics data. *BMC Bioinformatics* 2012; **13**(Suppl 16): S13.
- 3 Triska M, Grocutt D, Southern J, Murphy DJ, Tatarinova T. cisExpress: motif detection in DNA sequences. *Bioinformatics* 2013; **29**: 2203–2205.
- 4 Troukhan M, Tatarinova T, Bouck J, Flavell RB, Alexandrov NN. Genome-wide discovery of cis-elements in promoter sequences using gene expression. *OMICS* 2009; **13**: 139–151.
- 5 Mayer TZ, Simard FA, Cloutier A, Vardhan H, Dubois CM, McDonald PP. The p38-MSK1 signaling cascade influences cytokine production through CREB and C/EBP factors in human neutrophils. *J Immunol* 2013; **191**: 4299–4307.
- 6 Liu B, Sun L, Liu Q, Gong C, Yao Y, Lv X *et al*. A cytoplasmic NF-kappaB interacting long noncoding RNA blocks I kappa B phosphorylation and suppresses breast cancer metastasis. *Cancer Cell* 2015; **27**: 370–381.
- 7 Pallandre JR, Brillard E, Crehange G, Radlovic A, Remy-Martin JP, Saas P *et al*. Role of STAT3 in CD4+CD25+FOXP3+ regulatory lymphocyte generation: implications in graft-versus-host disease and antitumor immunity. *J Immunol* 2007; **179**: 7593–7604.
- 8 Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol* 2010; **11**: R106.
- 9 Benjamini Y, Heller R. Screening for partial conjunction hypotheses. *Biometrics* 2008; **64**: 1215–1222.
- 10 Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 2001; **125**: 279–284.
- 11 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 1995; **57**: 289–300.
- 12 Huang, da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009; **4**: 44–57.
- 13 Drees R, Ambrosini G, Cavin Perier R, Bucher P. EPD and EPDnew, high-quality promoter resources in the next-generation sequencing era. *Nucleic Acids Res* 2013; **41**: D157–D164.
- 14 Takahashi H, Kato S, Murata M, Carninci P. CAGE (cap analysis of gene expression): a protocol for the detection of promoter and transcriptional networks. *Methods Mol Biol* 2012; **786**: 181–200.
- 15 Maruyama K, Sugano S. Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* 1994; **138**: 171–174.
- 16 Tatarinova T, Kryshchenko A, Triska M, Hassan M, Murphy D, Neely M *et al*. NPEST: a nonparametric method and a database for transcription start site prediction. *Quant Biol* 2013; **1**: 261–271.
- 17 Caravaca JM, Donahue G, Becker JS, He X, Vinson C, Zaret KS. Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes. *Genes Dev* 2013; **27**: 251–260.
- 18 Ni M, Chen Y, Fei T, Li D, Lim E, Liu XS *et al*. Amplitude modulation of androgen signaling by c-MYC. *Genes Dev* 2013; **27**: 734–748.
- 19 Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010; **26**: 139–140.
- 20 Hodgson A, Wan F. Interference with nuclear factor kappaB signaling pathway by pathogen-encoded proteases: global and selective inhibition. *Mol Microbiol* 2016; **99**: 439–452.
- 21 Kulakovskiy IV, Medvedeva YA, Schaefer U, Kasianov AS, Vorontsov IE, Bajic VB *et al*. HOCOMO: a comprehensive collection of human transcription factor binding sites models. *Nucleic Acids Res* 2013; **41**: D195–D202.
- 22 Mathelier A, Zhao X, Zhang AW, Parcy F, Worsley-Hunt R, Arenillas DJ *et al*. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2014; **42**(Database issue): D142–D147.
- 23 Jolma A, Yan J, Whittington T, Toivonen J, Nitta KR, Rastas P *et al*. DNA-binding specificities of human transcription factors. *Cell* 2013; **152**: 327–339.
- 24 Kirsanov DD, Zanegina ON, Aksianov EA, Spirin SA, Karyagina AS, Alexeevskii AV. NPIDB: Nucleic acid-Protein Interaction DataBase. *Nucleic Acids Res* 2013; **41**: D517–D523.
- 25 Swindell WR, Sarkar MK, Stuart PE, Voorhees JJ, Elder JT, Johnston A *et al*. Psoriasis drug development and GWAS interpretation through in silico analysis of transcription factor binding sites. *Clin Transl Med* 2015; **4**: 13.
- 26 Vangala RK, Heiss-Neumann MS, Rangatia JS, Singh SM, Schoch C, Tenen DG *et al*. The myeloid master regulator transcription factor PU.1 is inactivated by AML1-ETO in t(8;21) myeloid leukemia. *Blood* 2003; **101**: 270–277.
- 27 Iwanaszko M, Kimmel M. NF-kappaB and IRF pathways: cross-regulation on target genes promoter level. *BMC Genomics* 2015; **16**: 307.
- 28 Lee PH, Puppi M, Schluns KS, Yu-Lee LY, Dong C, Lacorazza HD. The transcription factor E74-like factor 4 suppresses differentiation of proliferating CD4+ T cells to the Th17 lineage. *J Immunol* 2014; **192**: 178–188.
- 29 Zhang P, Chen HX, Duan YQ, Wang WZ, Zhang TZ, Li JW *et al*. Analysis of Th1/Th2 response pattern for erythrodermic psoriasis. *J Huazhong Univ Sci Technol Med Sci* 2014; **34**: 596–601.
- 30 Swindell WR, Johnston A, Carbajal S, Han G, Wohn C, Lu J *et al*. Genome-wide expression profiling of five mouse models identifies similarities and differences with human psoriasis. *PLoS ONE* 2011; **6**: e18266.
- 31 Lam EW, Brosens JJ, Gomes AR, Koo CY. Forkhead box proteins: tuning forks for transcriptional harmony. *Nat Rev Cancer* 2013; **13**: 482–495.
- 32 Massague J, Seoane J, Wotton D. Smad transcription factors. *Genes Dev* 2005; **19**: 2783–2810.
- 33 van Vliet J, Turner J, Crossley M. Human Kruppel-like factor 8: a CACCC-box binding protein that associates with CtBP and represses transcription. *Nucleic Acids Res* 2000; **28**: 1955–1962.
- 34 Chen ZY, Shie J, Tseng C. Up-regulation of gut-enriched kruppel-like factor by interferon-gamma in human colon carcinoma cells. *FEBS Lett* 2000; **477**: 67–72.
- 35 Guy LG, Delvoye N, Wall L. Expression of a human beta-globin transgene in mice with the CACC motif and upstream sequences deleted from the promoter still depends on erythroid Kruppel-like factor. *J Biol Chem* 2000; **275**: 3675–3680.
- 36 Ulgiati D, Subrata LS, Abraham LJ. The role of Sp family members, basic Kruppel-like factor, and E box factors in the basal and IFN-gamma regulated expression of the human complement C4 promoter. *J Immunol* 2000; **164**: 300–307.

- 37 Li B, Tsoi LC, Swindell WR, Gudjonsson JE, Tejasvi T, Johnston A *et al*. Transcriptome analysis of psoriasis in a large case-control sample: RNA-seq provides insights into disease mechanisms. *J Invest Dermatol* 2014; **134**: 1828–1838.
- 38 Kim KJ, Park S, Park YH, Ku SH, Cho EB, Park EJ *et al*. The expression and role of kruppel-like factor 4 in psoriasis. *Ann Dermatol* 2014; **26**: 675–680.
- 39 Keijsers RR, van der Velden HM, van Erp PE, de Boer-van Huizen RT, Joosten I, Koenen HJ *et al*. Balance of Treg vs. T-helper cells in the transition from symptomless to lesional psoriatic skin. *Br J Dermatol* 2013; **168**: 1294–1302.
- 40 Liu Y, Carlsson R, Comabella M, Wang J, Kosicki M, Carrion B *et al*. FoxA1 directs the lineage and immunosuppressive properties of a novel regulatory T cell population in EAE and MS. *Nat Med* 2014; **20**: 272–282.
- 41 DeGraff DJ, Grabowska MM, Case TC, Yu X, Herrick MK, Hayward WJ *et al*. FOXA1 deletion in luminal epithelium causes prostatic hyperplasia and alteration of differentiated phenotype. *Lab Invest* 2014; **94**: 726–739.
- 42 Grabowska MM, Elliott AD, DeGraff DJ, Anderson PD, Anumanthan G, Yamashita H *et al*. NF1 transcription factors interact with FOXA1 to regulate prostate-specific gene expression. *Mol Endocrinol* 2014; **28**: 949–964.
- 43 Albergaria A, Paredes J, Sousa B, Milanezi F, Carneiro V, Bastos J *et al*. Expression of FOXA1 and GATA-3 in breast cancer: the prognostic significance in hormone receptor-negative tumours. *Breast Cancer Res* 2009; **11**: R40.
- 44 Ye DZ, Kaestner KH. Foxa1 and Foxa2 control the differentiation of goblet and enteroendocrine L- and D-cells in mice. *Gastroenterology* 2009; **137**: 2052–2062.
- 45 Yoshimi T, Nakamura N, Shimada S, Iguchi K, Hashimoto F, Mochitate K *et al*. Homeobox B3, FoxA1 and FoxA2 interactions in epithelial lung cell differentiation of the multipotent M3E3/C3 cell line. *Eur J Cell Biol* 2005; **84**: 555–566.
- 46 Heddad Masson M, Poisson C, Guerardel A, Mamin A, Philippe J, Gosmain Y. Foxa1 and Foxa2 regulate alpha-cell differentiation, glucagon biosynthesis, and secretion. *Endocrinology* 2014; **155**: 3781–3792.
- 47 DeGraff DJ, Clark PE, Cates JM, Yamashita H, Robinson VL, Yu X *et al*. Loss of the urothelial differentiation marker FOXA1 is associated with high grade, late stage bladder cancer and increased tumor proliferation. *PLoS ONE* 2012; **7**: e36669.
- 48 Taube JH, Allton K, Duncan SA, Shen L, Barton MC. Foxa1 functions as a pioneer transcription factor at transposable elements to activate Afp during differentiation of embryonic stem cells. *J Biol Chem* 2010; **285**: 16135–16144.
- 49 Varley CL, Bacon EJ, Holder JC, Southgate J. FOXA1 and IRF-1 intermediary transcriptional regulators of PPARgamma-induced urothelial cytodifferentiation. *Cell Death Differ* 2009; **16**: 103–114.
- 50 Hurtado CW, Furuta GT, Kramer RE. Etiology of esophageal food impactions in children. *J Pediatr Gastroenterol Nutr* 2011; **52**: 43–46.
- 51 Kim TG, Lee HJ, Youn JI, Kim TY, Han H. The association of psoriasis with human leukocyte antigens in Korean population and the influence of age of onset and sex. *J Invest Dermatol* 2000; **114**: 309–313.
- 52 Romanowska M, al Yacoub N, Seidel H, Donandt S, Gerken H, Phillip S *et al*. PPARdelta enhances keratinocyte proliferation in psoriasis and induces heparin-binding EGF-like growth factor. *J Invest Dermatol* 2008; **128**: 110–124.
- 53 Swindell WR, Remmer HA, Sarkar MK, Xing X, Barnes DH, Wolterink L *et al*. Proteogenomic analysis of psoriasis reveals discordant and concordant changes in mRNA and protein abundance. *Genome Med* 2015; **7**: 86.
- 54 Li J, Li X, Hou R, Liu R, Zhao X, Dong F *et al*. Psoriatic T cells reduce epidermal turnover time and affect cell proliferation contributed from differential gene expression. *J Dermatol* 2015; **42**: 874–880.
- 55 Yoon S, Xuan Z, Makarov V, Ye K, Sebat J. Sensitive and accurate detection of copy number variants using read depth of coverage. *Genome Res* 2009; **19**: 1586–1592.
- 56 Yin X, Low HQ, Wang L, Li Y, Ellinghaus E, Han J *et al*. Genome-wide meta-analysis identifies multiple novel associations and ethnic heterogeneity of psoriasis susceptibility. *Nat Commun* 2015; **6**: 6916.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Supplementary Information accompanies the paper on Experimental & Molecular Medicine website (<http://www.nature.com/emm>)