

## Appendix S1

## SUPPLEMENTARY MATERIALS AND METHODS

*Double skinfold measurement*

The overall macroscopic effect of PUVA and UVB irradiation was assessed by measuring the double skinfold thickness (DSFT) of dorsal skin of the mice with a spring-loaded engineer's micrometer (Mitutoyo, Kanagawa, Japan) before the start of the experiment, on specific days during the experiment, and on the day the mice were sacrificed. Skin swelling (S) was determined by subtracting the skin thickness of the control group (Group VIII) from that of the treated group. The percentage suppression of skin swelling for a specific phototherapy group (PUVA or UVB) was determined using the following formula:  $(1 - (S_{\text{IMO+PUVA or UVB}} / S_{\text{IMO}})) \times 100$ . In addition, to quantify the response of the psoriatic dorsal skin to PUVA and UVB treatment a scoring system similar to the human Psoriasis Area and Severity Index (PASI) was used. In brief, erythema, infiltration and scaling of the skin were scored "blind" on a scale from 0 to 3, as follows: 0: none; 1: slight; 2: moderate; 3: severe. The single scores were summed, resulting in a theoretical maximal total score of 9.

*Tissue collection*

Mice were sacrificed at the time-point indicated in Fig. S1<sup>1</sup> and samples of dorsal skin, spleen, lymph nodes, and blood were collected. Approximately 1 cm<sup>2</sup> of central dorsal skin per mouse was excised, fixed immediately in 4% buffered formaldehyde, processed routinely, and sectioned at 4 µm for haematoxylin and eosin (H&E) staining. In addition, tissue was submerged in RNA later solution (Applied Biosystems, Foster City, CA, USA) and stored at -70°C for later mRNA analysis.

*Histological evaluation*

Epidermal hyperplasia was assessed by counting epidermal cell layers and measuring the thickness of the epidermis from the basal layer to the stratum. For quantification of epidermal thickness and layers, 5 randomly selected locations per H&E-stained cross-section of dorsal skin from each mouse were examined under the microscope at 20× magnification (scale bar 100 µm). All measurements were performed blind. Results were first averaged per mouse and then per treatment group in the statistical analyses. The percentage suppression in the increase in epidermal layers and skin thickness for specific groups was determined in an analogous way to that of suppression of skin swelling, using the formula given above. Representative images were acquired with a DP71 digital camera (Olympus, Melville, NY, USA) attached to an Olympus BX51 microscope.

*Immunohistochemistry*

Paraffin-embedded sections were stained for p21, p16 and phospho-Histone H2A.X (Ser139). De-waxed and rehydrated tissue sections were heated in 0.01 M citrate buffer at pH 6.0 in an autoclave, followed by an endogenous peroxidase activity block in 3% hydrogen peroxide in phosphate-buffered saline (PBS). Blocking steps were then performed using the Avidin/Biotin blocking kit, SP2001 (Vector Laboratories, Burlingame, CA, USA), super block, IDSTM003 (Empire Genomics, New York, NY, USA) and mouse block (Empire Genomics, IDSTM003). The blocked sections were incubated with the primary antibody p21 (Santa Cruz; sc-397) 1:100 dilution, p16ink4a, ab54210 (Abcam, Cambridge, UK) 1:50 dilution, and phospho-Histone H2A.X (Ser139), #9718 (Cell Signalling, Leiden, The Netherlands) 1:100 dilution at 4°C overnight. The detection was done with the I Detect Super Stain

System HRP (Empire Genomics, IDSTM003) and the specific signal was visualized with 3-amino-9-ethylcarbazole, BP1108 (ID Laboratories, New York, NY, USA) followed by counterstaining with haematoxylin.

*Immunoscore*

Membranous and cytoplasmic/endogenous levels of protein immunoreactivity for p21 and p16 protein were analysed and graded by the percentage of positive cells within the epidermis.

*Multiplex immunofluorescent analysis*

The Opal™ 6-Color Fluorescent IHC Kit (Perkin Elmer, Vienna, Austria) was used according to the manufacturer's instructions. Slides were stained with antibody specific for dendritic cells maturation marker, MHC-II (Abcam). Abcam CD3 antibody (Abcam) and Biorad Ly6B antibody (Hercules, CA, USA) antibodies were also used together with MHC-II to achieve better separation of the MHC-II positive population. For image acquisition, the Vectra® 3 automated quantitative pathology imaging system (PerkinElmer) was used to detect the relative abundance of cells expressing the marker protein in the epidermis and dermis upon tissue segmentation.

*Bead immunoassay*

Procarta 26-Plex bead immunoassay from Affymetrix eBioscience (EPX260-26088-901) for detecting IFN-γ, IL-12p70, IL-13, IL-1 beta, IL-2, IL-4, IL-5, IL-6, TNF alpha, GM-CSF, IL-18, IL-10, IL-17A, IL-22, IL-23, IL-27, IL-9, GRO alpha, IP-10, MCP-1, MCP-3, MIP-1 alpha, MIP-1 beta, MIP-2, RANTES and eotaxin was used according to the manufacturer's specifications. Analysis was performed with 5 parametric curve fittings (Bio-Rad, Hercules). Standard curves for each analyte were generated by using the reference analyte concentration supplied by the manufacturers.

*RNA isolation*

Total RNA from frozen dorsal skin samples was extracted with the miRNeasy Kit (Qiagen, Hilden, Germany; catalogue number 217004) including DNase treatment steps on the column according to protocol and homogenized with MagNA Lyser Green Beads (Roche, Basel, Switzerland; catalogue number 03358-941-001) using the MagNA Lyser (Roche Basel, Switzerland). RNA was quantified by measuring absorbance at 260 and 280 nm, and quality was secured with a ratio between RNA and protein of 1.9–2.09 using a Nano Drop (Thermo Fisher Scientific, Waltham, MA, USA). Furthermore, degradation was checked on the Bio-Analyzer BA2100 (Agilent, Foster City, CA, USA; catalogue number 5065-4476).

*RNA microarray and pathway analysis*

GeneChip® Mouse Gene 2.0 ST Arrays (Affymetrix, Santa Clara, CA, USA; catalogue number 902118) were used for whole transcript amplification in comparison with 3'IVT amplification, which depends on best-quality RNA. 500 ng of total RNA was used for the amplification, following the manufacturer's protocol. The amplified cDNA was analysed again using the BioAnalyzer BA2100 (Agilent, Foster City, CA, USA) with the RNA 6000 Nano LabChip (Agilent, Foster City, CA, USA; catalogue number 5065-4476). The given fragment size of <2000 nt for overall samples passed our quality control to proceed with the ss-cDNA synthesis, fragmentation and labelling. These were hybridized for 18 h at 45°C, according to the supplier's protocol, while rotating in the hybridization oven. Washing and staining (GeneChip® HT hybri-

dization, Wash and Stain Kit; Affymetrix, Santa Clara, CA, USA; catalogue number 900720) were performed with the Affymetrix Genechip® fluidics station 450 according to the manufacturer's protocol (protocol on fluidics station: FS450\_0002). Arrays were scanned with the Affymetrix GeneChip scanner GCS3000. To evaluate the hybridization controls and pre-analysis, Affymetrix Expression Console EC 1.3.1 was used with no technical outlier arrays being detected. Raw data are available at the Gene Expression Omnibus with GEO; with accession number GSE100774. Statistical analysis for this data set was performed with the Partek Software v.6.6 (Partek Inc., St Louis, MO, USA). CEL files with the probe intensity data were imported using the robust multi-chip average (RMA) algorithm. This included background correction, quantile normalization across all arrays, and median polished summarization based on log-transformed expression values. Differences among groups were tested using 1-way ANOVA. For pathway analysis 194 transcripts were used, showing a sig-

nificant  $p$ -value  $<0.05$  and a fold change  $\pm 1.5$  between Groups V and VI. Ingenuity Pathway Analysis software (Qiagen) was used for pathway analysis performed with the following settings: Species=Mouse AND (confidence=Experimentally Observed OR High (predicted)).

#### *Statistical analysis*

Each experiment was repeated at least once with similar results. Data presented are expressed as means  $\pm$  standard error of the mean (SEM). Statistical differences among experimental groups were determined by use of ANOVA or paired or unpaired  $t$ -test or Mann-Whitney  $U$  test after testing for normality, whichever was deemed appropriate. The statistical test chosen for each experiment is indicated in the figure legend. Statistical significance was set at  $p \leq 0.05$  and levels of significance were given as \* $p \leq 0.05$ ; \*\* $p \leq 0.01$ ; \*\*\* $p \leq 0.001$ ; and \*\*\*\* $p \leq 0.0001$ .