Appendix S1

MATERIALS AND METHODS

The human biological samples were sourced ethically and their research use was in accord with the terms of the informed consents. This study was approved by the ethics committees of all institutions involved.

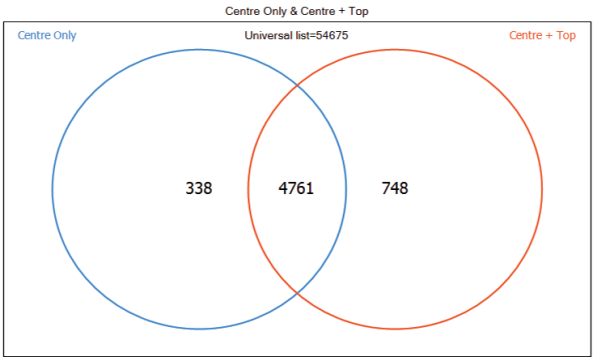
Microarray meta-analysis and identification of DEGs

For the whole keloid biopsy study, 5 pre-sternal keloid biopsies and 5 normal skin biopsies were harvested from a total of 10 Caucasian subjects (5 subjects who suffered from KD, 5 subjects who did not present with KD, respectively) as previously published (31). The keloid biopsies were sectioned by region into leading edge, centre, or top areas (see Fig. 1). Amplified RNA isolated from each sample was hybridised onto Agilent Whole Human Genome (4×44K) Oligo Microarrays (Agilent, Santa Clara, CA, USA). The microarray data for a leading edge biopsy (keloid donor 103) was discarded due to poor quality. Agilent text files were imported into Array Studio v6.0 (OmicSoft, Cary, NC, USA) using the default parameters and Median Absolute Deviation (MAD) was performed to identify outliers (outlier defined as having absolute MAD score >5). 19 samples were taken forward for statistical analysis and the following comparisons were estimated using t-tests: keloid leading edge ($n=4$) versus normal skin ($n=5$); keloid centre ($n=5$) versus normal skin ($n=5$); keloid top ($n=5$) versus normal skin ($n=5$). The list of DEGs was filtered to remove low intensity probe sets (detected probe sets were defined as having log2 maximum median group intensity >7, leaving 23599) and were loaded into IPA (Ingenuity Systems Inc., Redwood City, CA, USA). Data was subjected to IPA interrogation on probe sets that showed a 1.2-fold change with $p<0.05$.

For the *ex vivo*-cultured dermal fibroblast microarray study, 3 healthy Caucasians with spontaneous KD with the presence of keloids for at least 1 year were enrolled in the study as previously described 30). The keloids were excised and subcutaneous tissue was carefully removed and the edges of the keloids were trimmed to remove unaffected tissue. Five millimetre punch biopsies were then taken from the leading edge, centre, or top areas (see Fig. 1). 5-mm punch biopsies were taken from the healthy skin of 3 Caucasian volunteers who did not present with KD. Dermal fibroblasts were isolated and cultured as previously described and RNA from each sample was isolated and quantified, and equal amounts from each keloid site and from the healthy skin were pooled into 4 single samples. RNA was amplified using standard Affymetrix protocols, and for each pool of biotin-labelled RNA, 2 Affymetrix HG-U133_Plus2.0 microarrays were hybridised (Affymetrix, Santa Clara, CA, USA). One Sample from the keloid top (LJ0015) was not hybridised due to poor RNA quality. The final numbers of samples were: normal skin ($n=2$), keloid leading edge ($n=2$), keloid centre ($n=2$) and keloid top ($n=1$) (30). CEL files were loaded into Array Studio v6.0, GCRMA normalised and MAS5.0 QC metrics were generated. 7 samples were taken forward for statistical analysis and the following comparisons were estimated using *t*-tests: keloid leading edge ($n=2$) versus normal skin ($n=2$); keloid centre ($n=2$) and keloid top ($n=1$) versus normal skin ($n=2$). Principal component analysis on the detected probe sets indicated that the cultured fibroblasts from the top of the keloid and the centre of the keloid were very similar (Fig. 2A), but distinct from the fibroblasts cultured from the leading edge of the keloid. Therefore at the gene expression level, there would be little value in treating the keloid top fibroblasts separately, as there was only $n=1$ for this group. By combining the data from the keloid top fibroblasts with keloid fibroblasts from the centre, this sample can be included in the final analysis. A sensitivity analysis to understand the effect of combining the top and centre keloid fibroblasts was performed by repeating the ANOVA including only the centre samples. The Venn diagram in SFig. 1 shows that the overlap between the original analysis (centre + top, right side of Venn) and the sensitivity analysis (centre only, left side of Venn) is very high with between 80–90% of significant probe sets appearing in both lists. Of the 748 probe sets that are significant in the Centre + Top (original analysis, $n=3$) only, most (>50%) of these are just missing significant ($p<0.1$) in the Centre only (sensitivity analysis, $n=2$), indicating this is likely due to a lack of statistical power with $n=2$ rather than differences in expressed genes. The Centre + Top list was taken forward for the final fibroblast analysis. The list of DEGs was filtered to remove low intensity probe sets (detected probe sets were defined as having log2 maximum median group intensity >4, leaving 13947). Data was subjected to IPA on probe sets that showed a 1.2-fold change with $p<0.05$.

Measurements of interleukin 11 secretion from TGFβ1-stimulated dermal fibroblasts

Dermal fibroblasts from 5 healthy donors (Clonetics, Lonza, Rockland, ME, USA), were cultured according to the manufacturer's guidelines. Briefly, cells were cultured in 0.4% (v/v) FCS/DMEM were un-stimulated (vehicle-treated) or stimulated with 0.5 ng/ml human TGFβ1 at 37ºC for 24 h. IL-11 secretion was used as a positive control for response to TGFβ1 assessed using IL-11 MSD assay kits (Meso Scale Discovery, Rockville, MD, USA). Coat standard bind plate with capture ab at 0.5 µg/ml in PBS. Incubate at +4C o/n, wash plate 3x with 1x MSD Tris wash buffer, add blocker, 150 µl/well and incubate for 1 h at RT on plate shaker (600 rpm), prepare standard curve of IL11 in matrix buffer, e.g. 2,000 to 15.6 pg/ml 1/2 dilution series, defrost cell supernatants to be tested, wash plate 3×with 1× MSD Tris wash buffer, add standard curve dilutions in duplicate 25 µl/well, add samples to plate in duplicate 25µl/well, incubate for 2 h at RT on plate shaker (600 rpm), wash plate 3× with 1× MSD Tris wash buffer, add detection ab at 1µg/ml in assay diluent. Incubate for 1 h at RT on plate shaker (600 rpm), wash plate 3× with 1× MSD Tris wash buffer, add 2× Read Buffer T 150 µl/well and read on MSD plate reader, analyse on MSD Software, xlfit or GraphPad Prism.



SFig. 1. Venn diagram showing overlap of sensitivity analysis to assess effect of combining fibroblasts isolated from the top and centre keloid regions. A sensitivity analysis was performed to demonstrate that combining the fibroblasts isolated from the centre and top of keloids would not have a large impact on the final list of differentially expressed genes. The Venn diagram above shows that the overlap between the original analysis (centre + top, right side of Venn) and the sensitivity analysis (centre only, left side of Venn). The overlap between both lists of differentially expressed genes is very high (80–90%), and of the 748 probe sets that are significant in the Centre + Top (original analysis, $n=3$) only, most (>50%) of these are just missing significant ($p<0.1$) in the Centre only (sensitivity analysis, $n=2$), indicating this is likely due to a lack of statistical power with $n=2$ rather than differences in expressed genes.

Identification of DEGs following TGFβ stimulation of dermal fibroblasts and robust keloid signatures from different spatial locations

Dermal fibroblasts from 5 healthy donors (Clonetics, Lonza, Rockland, ME, USA), were cultured according to the manufacturer's guidelines. Dermal fibroblasts grown in 0.4% (v/v) FCS/DMEM were un-stimulated (vehicle-treated) or stimulated with 0.5 ng/ml human TGFβ1 (R&D Systems, Abingdon, UK) at 37°C for 8 h. RNA isolation was conducted using the QIAshredder protocol (QIAGEN, Venlo, Netherlands). Purified total RNA was amplified and labelled using Nugen Ovation kits (Nugen Technologies Inc., San Carlos, CA, USA). Samples were hybridized to Affymetrix HG-U133_plus_2.0 microarrays and arrays were washed, stained and scanned using standard Affymetrix procedures. CEL files were loaded into Array Studio v6 and normalised using GCRMA and MAS5.0 QC metrics were generated. A linear model factoring group was applied (where group has levels: un-stimulated, TGFβ1-stimulated) and comparisons between groups conducted. DEGs were estimated using student's *t*-test and low intensity probe sets were filtered (detected probe sets were defined as having log2 maximum median group intensity >5, leaving 21167).

As the keloid biopsy and keloid fibroblast data sets were profiled on different microarray platforms (keloid fibroblasts on Affymetrix HG-U133_plus_2.0, keloid biopsy on Agilent 4×44K), a robust signature was defined at the gene level rather than at the probe set level due to differences in probe set sequence and length. Each gene list was filtered for detected probe sets (see criteria above), and collapsed to the gene level, where a gene was represented by multiple probe sets a filter for significance was first applied ($p<0.05$), and if multiple probe sets were significant, then the probe set with the largest fold-change was selected. For the leading edge and centre gene lists from each study, a robust list was defined as any gene that showed a significant ($p<0.05$) change in both the keloid biopsy and keloid fibroblast data sets in the same direction with a magnitude of at least 1.2-fold. For the keloid top signature (keloid biopsy dataset only), a signature was defined as a significant change ($p<0.01$) with a fold-change of at least 2-fold in either direction. The TGFβ1-stimulated dermal fibroblast gene signature was collapsed to the gene level using the same approach.

Robust signatures were loaded into IPA using the search function and the connect function was used to determine whether any known interactions had previously been published.

GeneSet Enrichment Analysis

A universe was defined as all genes that were represented on both the Affymetrix HG-U133_plus_2.0 and the Agilent 4x44K microarrays, leaving 17,993 unique genes. The TGFβ1-stimulated dermal fibroblast gene signature and robust keloid gene signatures were filtered for genes in the defined universe. The lists were then separated into up- and down-regulated genes and a 4-way Venn diagram was generated comparing the up- and down-regulated genes from the TGFβ1-stimulated dermal fibroblasts gene signature to the up- and down-regulated genes from the keloid leading edge, keloid centre and keloid top gene signatures. The number of genes changing in the same direction between the TGFβ1-stimulated dermal fibroblast gene signature and the keloid signatures was calculated and tested for statistical significance using Fisher's Exact Test.

Gene Interaction analysis

In order to help understand and interpret the lists of differentially expressed genes identified from the individual data sets analysed, and the combined analysis multiple interaction analyses were performed in Ingenuity Pathway Analysis.

Differentially expressed gene lists from the individual study analyses were loaded in to Ingenuity Pathway Analysis and upstream regulator analysis was performed. Ingenuity's Upstream Regulator Analysis is a unique tool that both predicts upstream regulators from gene expression data and whether each regulator is in an activated or inhibited state. An upstream regulator is defined as any gene, RNA, protein, or chemical that can affect the expression of a gene. Two separate metrics are calculated and can be used together or independently for biological interpretation. The first metric, an activation Z-score, is based on the consistency of literature-derived expression effects between an upstream regulator and its gene targets, and the measured differential expression; the expected expression effects are derived from the literature and compiled in the Ingenuity® Knowledge Base. The Z-score indicates the degree of consistent agreement or disagreement of the actual versus the expected direction of change among downstream gene targets. A prediction about the state of the upstream regulator, either activated or inhibited, is made based on the Z-score.

For example, if most of the targets of an upstream regulator are expected from the literature to be upregulated and there is an observed increase in their measured gene expression in the analyzed dataset, this would lead to a positive Z-score and an "activation" prediction. Conversely, if most of the genes were expected to be upregulated and the observed expression was down-regulated (or vice-versa), this anti-correlation would lead to a negative Z-score and a prediction of "inactivated" for the upstream regulator. Therefore, the larger the absolute Z-score, the stronger the evidence for the role of the upstream regulator.

As an independent calculation, regardless of gene expression direction change, a Fisher's Exact Test *p*-value is calculated to assess the significance of enrichment between downstream genes of an upstream regulator and those observed in the gene list. Note this *p*-value does not take direction into account. Genes identified as robust gene signatures for the distinct keloid regions were subject to interaction analysis in Ingenuity Pathway Analysis. In order to complete this analysis, the list of robust genes from each signature was pasted into Ingenuity Pathway Analysis search function and mapped genes were loaded into a new network. The connect function was then used to identify interactions between the genes. Subsets of genes from each list that formed distinct networks were identified and have been reported in Fig. 3, Figs S1 and S2[1]. Note in Figs S1 and S2[1], then networks have been trimmed to focus on the central nodes (genes with multiple connections) to allow easier interpretation.

GSEA in psoriasis and atopic dermatitis

A similar approach was used to understand the role of TGFβ1 in other dermal conditions. Using published microarray data sets, a gene signature was defined for psoriasis (S1) and atopic dermatitis (S2–S4). SFig. 2 shows the overlap between the TGFβ1 signature and the psoriasis signature (A) and the atopic dermatitis signature (B). The results from the Fisher's Exact Test indicated that there was no significant overlap between the TGFβ1 signature and the psoriasis signature (P=1), but that there was a significant overlap between the TGFβ1 signature and the atopic dermatitis signature ($p=2.1\times10^{-11}$).

Real Time PCR

RNA from 3 healthy skin samples and 11 keloid biopsies (from 7 patients, keloids were either untreated or saline injected) was isolated using RNeasy universal Plus mini kit as per manufacturers instructions. RNA QC was performed using the Bioanalyzer 2100 (Agilent Technologies) and quantified using a Nanodrop (Thermo Scientific).

SFig 2. Venn diagrams showing the overlap between TGFβ-1 stimulated dermal fibroblasts and psoriasis (A) and atopic dermatitis (B). To understand the robustness of the overlap between the keloid signatures and TGFβ-1, a similar analysis was performed using gene signatures from psoriasis and atopic dermatitis. The microarray data sets used for this analysis were obtained through the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo/). For psoriasis GSE13355 was used, and for atopic dermatitis GSE6012 and GSE16161 were used to generate disease signatures. When the overlap was tested for significance using Fisher's Exact Test, the overlap for psoriasis was not significant ($p = 1$) whereas the overlap for atopic dermatitis was highly significant ($p = 2.1 \times 10^{-11}$).

1 µg of total RNA was reversed transcribed into cDNA using Invitrogen's Superscript Vilo cDNA synthesis kit (cat 11754-050) in a 20 µl reaction. Taqman qPCR reactions were run in triplicate using 10ng cDNA per reaction (Taqman universal gene expression mastermix, Applied Biosystems cat 4304437). Applied Biosystems off the shelf taqman assays were used at stock concentrations. Taqman cycling parameters were 50ºC for 2 min, 95ºC for 10 min, followed by × 40 cycles of 95ºC for 15 s and 60ºC for 1 min.

GAPDH, ACTB and *PPIA* were run as housekeepers and *POSTN, COMP, ADAM12, COL11A1, ITGA10, OPCML* and *SYNDIG1 (TMEM90B)* were run as test genes. CT data was normalised using the scores from the first component from a principal component analysis on using just the housekeeper genes. Normalised CT values were then converted to abundance using the following equation: abundance = 10^((40-CT)/3.5). Log2 abundances were used to calculate differential expression (*t*-test, Benjamini-Hochberg false discovery rate, FDR) and plotted using R gplot boxplot (SFig. 3).

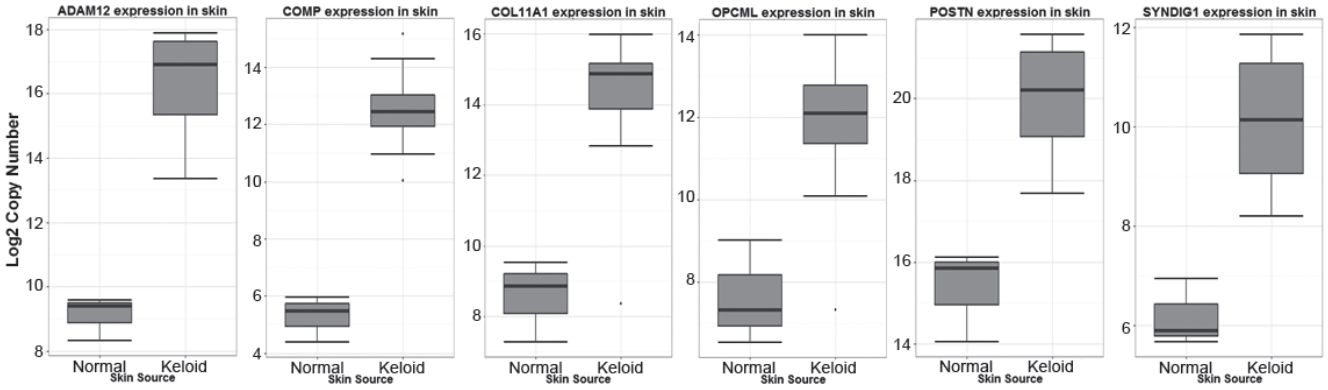
SUPPLEMENTARY REFERENCES

S1. Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D, et al. Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. Nat Genet 2009; 41: 199–204.

S2. Olsson M, Broberg A, Jernas M, Carlsson L, Rudemo M, Suurkula M, et al. Increased expression of aquaporin 3 in atopic eczema. Allergy 2006; 61: 1132–1137.

S3. Plager DA, Leontovich AA, Henke SA, Davis MD, McEvoy MT, Sciallis GF, 2nd, et al. Early cutaneous gene transcription changes in adult atopic dermatitis and potential clinical implications. Exp Dermatol 2007; 16: 28–36.

S4. Guttman-Yassky E, Suarez-Farinas M, Chiricozzi A, Nograles KE, Shemer A, Fuentes-Duculan J, et al. Broad defects in epidermal cornification in atopic dermatitis identified through genomic analysis. J Allergy Clin Immunol 2009; 124: 1235–1244 e58.



SFig. 3. Real time PCR validation in independent keloid tissue. All fold changes were calculated using a *t*-test on log2 abundance data. *p*-value was Benjamini-Hochberg corrected. *ADAM12* 157.3 fold up-regulated $p < 0.0001$ in keloid; *COMP* 151.9 $p < 0.0001$ up-regulated in keloid; *COL11A1* 45.5 fold $p = 0.0021$ up-regulated in keloid; *OPCML* 17.4 fold $p = 0.0053$ up-regulated in keloid; *POSTN* 25.7 fold $p = 0.0003$ up-regulated in keloid; *SYNDIG1* 15.8 fold $p = 0.0006$ up-regulated in keloid. Data was plotted in R using ggplot2 boxplot; the median is shown and the box represents the first and third quartiles, whiskers represent the highest/lowest value that is within 1.5 * inter-quartile range of the box. Data beyond the end of the whiskers are outliers and plotted as points (as specified by Tukey).