An active chromatin interactome elucidates the biological mechanisms underlying genetic risk factors of dermatological conditions in disease relevant cell lines

Chenfu Shi[1*], Helen Ray-Jones[1,2*], James Ding[1], Kate Duffus[1], Yao Fu[3], Vasanthi Priyadarshini Gaddi[1], Oliver Gough[1], Jenny Hankinson[4], Paul Martin[1,5], Amanda McGovern[1], Annie Yarwood[1,2], Patrick Gaffney[3], Steve Eyre[1], Magnus Rattray[6], Richard B Warren[2] and Gisela Orozco[1].

1) Centre for Genetics and Genomics Versus Arthritis. Division of Musculoskeletal and Dermatological Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, United Kingdom
2) Dermatology Centre, Salford Royal NHS Foundation Trust, Manchester NIHR Biomedical Research Centre, Manchester Academic Health Science Centre, Manchester, United Kingdom
3) Genes and Human Disease Research Program, Oklahoma Medical Research Foundation, Oklahoma City, 73104, OK, USA
4) Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, University of Manchester, Manchester, UK
5) The Lydia Becker Institute of Immunology and Inflammation, University of Manchester, Manchester, UK
6) Division of Informatics, Imaging and Data Sciences, Faculty of Biology, Medicine and Health, University of Manchester, UK
*These authors contributed equally to this manuscript

Abstract

In the last 15 years Genome wide association studies (GWAS) have uncovered the genetic factors that contribute to disease risk for many disorders, yet the biological significance of many of these associations remain unclear. This is because about 90% of these variants are predicted to affect regulatory elements that are highly cell-type specific and can affect distal genes through chromatin interaction mechanisms and, therefore, understanding their role in disease is challenging.

Previous studies have attempted to use chromatin conformation methods to describe the functional mechanism in disease associated loci. However, the majority of these studies have focused on cell lines derived from blood, which might not be the most relevant cell lines for dermatological conditions.

Here, we generated HiChIP for H3K27ac, Hi-C and RNA-seq datasets from a keratinocyte and a skin-plaque derived CD8+ T cell line. We studied their active chromatin conformation with the aim to identify biological mechanisms that could be affected by variants associated with skin related diseases: psoriasis (Ps), psoriatic arthritis (PsA), systemic sclerosis (SSc), melanoma and atopic dermatitis.

In our analysis we show that HiChIP interactions provide functional evidence for gene regulation by showing that enhancers linked to genes that respond to IFN-γ stimulation are strongly enriched for motifs for TFs related to the stimulation. We also show that by using chromatin conformation we can recall 53% of eQTLs identified by GTEx in skin.

In addition to our datasets, we integrated public datasets for a B cell-like lymphoblastoid cell line and primary CD4+ T cells to expand our analysis to include GWAS variants associated with the aforementioned diseases likely to be mediated specifically by B cells and CD4+ T cells (such as in systemic sclerosis and melanoma). We then identified the numerous genes that interact with GWAS

variants associated and show that these genes strongly enrich for pathways that are related to the trait studied. For example, the top enriched pathways for autoimmune conditions such as PsA and psoriasis included responses to cytokines and T cell regulation, whereas for melanoma they were related to replicative senescence.

We show examples of how our analysis can inform changes in the current description of a few psoriasis associated risk loci. For example, the variant rs10794648, which had previously been assigned to IFNLR1, was linked to GRHL3 in our dataset, a gene essential in skin repair and development. This can indicate a renewed importance of skin related factors in the risk of disease.

In conclusion, this study allows us to further characterize disease risk loci by using novel chromatin conformation techniques, which can help identify the genes that are putatively involved in disease and has the potential to identify therapeutic targets.

Introduction

In the last 15 years GWAS studies have uncovered the genetic factors that contribute to disease risk for many complex disorders. It is now accepted that the majority of these genetic risk factors do not influence directly coding sequences but rather regulatory regions such as enhancers and promoters which can be highly cell type specific (Ernst *et al.* 2011; Farh *et al.* 2015; Kundaje *et al.* 2015). Many studies have also demonstrated that the targets of these variants cannot be assigned to the gene that is closest to them as they can affect genes that are very far away from their genomic location via chromatin looping mechanisms (GTEx Consortium 2013; Nica and Dermitzakis 2013; Rao *et al.* 2014; Javierre *et al.* 2016; Aguet *et al.* 2017; Võsa *et al.* 2018).

Recently, there has been a growing interest in using chromatin conformation and other functional genomics techniques to describe these disease-associated loci and identify the genes that are affected by them. Previous studies have used techniques such as Capture Hi-C and HiChIP to link the genes that physically interact with disease associated loci (Dryden *et al.* 2014; Jäger *et al.* 2015; Martin *et al.* 2015, 2016; Cairns *et al.* 2016; McGovern *et al.* 2016; Mumbach *et al.* 2017).

These studies have mainly focused on cells derived from blood immune cells. Multiple publications have shown that interactions are cell type specific and are altered during differentiation and stimulation (Dixon *et al.* 2012; Rao *et al.* 2014; Schmitt *et al.* 2016; Burren *et al.* 2017; Mumbach *et al.* 2017; Rubin *et al.* 2017; Siersbæk *et al.* 2017; Hansen *et al.* 2018). Although for many autoimmune diseases these genetic factors primarily affect immune cells, other cell types might be involved in the development of these diseases (Farh *et al.* 2015; Mahajan *et al.* 2018; Mizoguchi *et al.* 2018).

Skin disorders such as psoriasis, psoriatic arthritis and systemic sclerosis can involve cell types such as keratinocytes, the most predominant cell type in the epidermis and often highly dysregulated in disease (Mccoy *et al.*; Albanesi *et al.* 2018). For many of these disorders a recurrent feature is also the invasion of immune cells such as $CD8^+$ T cells in the inflammation site (O'reilly, Hü Gle and Van Laar; Hennino *et al.* 2007; Cai, Fleming and Yan 2012; Antohe *et al.* 2019). These cells can be altered and can have different phenotypes compared to circulating/naïve cells. For these diseases with high skin involvement, studying interactions from these cell types would be valuable in advancing the study of the mechanism of disease, but this has been limited by the cost required to fully map genome-wide interactions across all cell types.

HiChIP is a recently developed technique that allows highly sensitive detection of long range interactions at lower cost by focusing the sequencing efforts in specific regions of the genome (Mumbach *et al.* 2016, 2017). This is done by using chromatin immunoprecipitation to capture genomic regions associated with a specific histone modification of interest following the Hi-C library preparation. By focusing the sequencing efforts on these regions, it allows higher sensitivity and resolution compared to Hi-C.

In our previous work, we used region capture Hi-C to identify chromatin interactions at regions of the genome that contain variants associated with psoriasis in keratinocytes and skin derived $CD8^+$ T cells (Ray-Jones *et al.* 2019). This targeted study suggested novel candidate causal genes for known psoriasis risk loci, such as the *KLF4* gene at the intergenic locus 9q31.

Here we use HiChIP to map active chromatin interactions genome wide on keratinocytes and skin derived CD8+ T cells. We show that this technique, in contrast with region capture Hi-C, shows significantly better enrichment for active regions of the genome. Moreover, it analyses interactions genome-wide, allowing us to discover candidate genes for a larger set of disorders and including

more recently identified loci in our analysis. In addition to naïve cells we have also generated data from interferon gamma (IFN-γ) stimulated keratinocytes. IFN-γ is an inflammatory mediator implicated in multiple immune-mediated conditions such as systemic sclerosis (Wu and Assassi 2013), psoriasis and others (Lees 2015). In psoriasis in particular, psoriatic lesions are highly enriched in IFN-γ (Uyemura *et al.* 1993; Schlaak *et al.* 1994). IFN-γ promotes epidermal keratinocyte apoptosis (Hijnen *et al.* 2013) and represents the classical Th1 pathway whereby IFN-γ is released in abundance by activated T cells in psoriatic skin (Lowes *et al.* 2008).

We complement our dataset with matched RNA-seq and Hi-C datasets to increase the information gained from these cell types. Because the disorders studied have a significant immune component we include in our analysis publicly available HiChIP data for Naïve $CD4^+$ T cells and a B cell-like lymphoblastoid cell line (Mumbach *et al.* 2017) to describe the GWAS associated loci which might be mediated by these cell populations.

After identifying the genes that were linked by chromatin interaction to GWAS loci we show that that these genes enrich for pathways that are highly relevant to the underlying disease mechanisms. We show that using the results generated by our analysis that we can provide a functional mechanism that drive disease susceptibility in some loci. More importantly, we present how our results can also inform the association of different genes to specific loci and provide example of these for 4 distinct psoriasis loci. This has the impact of updating our view on the underlying disease mechanisms of these traits which can have a major impact in future studies and drug discovery.

Results

A compendium of activity and chromatin interactions in keratinocytes.

In this study we present a series of functional genomics datasets that we have produced for two cell lines that are relevant in a number of disorders involving skin, which we then applied to uncover novel functional information about these disorders. We used a number of techniques including Hi-C and HiChIP for H3K27ac to study the three-dimensional conformation of the genome and complemented it with RNA-seq to study the transcriptome and applied it to it to two cell lines: HaCaT cells, which are spontaneously immortalised keratinocytes, and MyLa cells, which are CD8+ T cells extracted from a cancerous skin plaque. We complemented our analysis with HiChIP data from the public domain for naïve T cells and B cell line GM12878 (Mumbach *et al.* 2017) to describe the disease associated loci likely mediated by these cell types.

We used HiChIP to identify interactions that are specific for active regulatory elements such as enhancers and promoters, through chromatin immunoprecipitation for the active histone mark H3K27ac. We analysed H3K27ac HiChIP data using a recently developed software specific for HiChIP data, FitHiChIP (Bhattacharyya *et al.* 2019). This software identifies significantly enriched interactions with a much higher accuracy than previously available software. Interactions and peaks are highly reproducible and cell type specific, as shown by unsupervised clustering and principal component analysis (figure 1A-C). Because the samples cluster strongly by condition, we decided to merge the two replicates together to increase power as we noticed that the number of significant interactions reported by FitHiChIP depends on the read depth. For example, in HaCaT naïve the two replicates yielded 13950 and 38118 interactions, while the merged dataset yielded 58268 interactions.

We identified more than 50000 significant interactions genome wide from each of our HiChIP datasets (summary statistics in table 1). The median interaction distance was 250 kb (representative distribution plotted in figure 1D). These values are consistent with the results derived from the public datasets. Moreover, the vast majority (90%) of interactions were within Topologically associating domains (TADs) identified from our matched Hi-C datasets (figure 2C). This is consistent with recent reports about TADs determining the scope of gene regulation (Rao *et al.* 2014; Rowley and Corces 2018).

Because many skin related conditions represented inflamed states, we stimulated HaCaT cells with INF-γ. INF-γ is linked to conditions such as psoriasis and the response that these cells show to INF-γ stimulation would inform stimulation specific interactions relevant in disease. Data from RNA-seq shows successful stimulation with 535 genes differentially expressed and clearly enriching for pathways related to INF gamma stimulation (Supplementary figure S1).

HiChIP strongly enriches for interactions involving active regions of the genome in contrast with capture Hi-C

We compared these results with previous region Capture Hi-C data which we generated for HaCaT and MyLa cells (Ray-Jones *et al.* 2019). These libraries specifically captured about 4500 HindIII fragments linked to GWAS SNPs for psoriasis, juvenile idiopathic arthritis, asthma, psoriatic arthritis, rheumatoid arthritis and systemic sclerosis, curated from the largest available meta-analyses at the time and additional smaller GWAS.

Our previous region capture Hi-C study identified about 35,000 significant interactions, with a median interaction distance of 280 kb (figure 2A). Similar to HiChIP data, the majority of the interactions reside within TADs (Figure 2C).

Comparing HiChIP with Capture Hi-C we see that Capture Hi-C interactions are not specifically selected for active genes or enhancers and, although significant interactions are still enriched for H3K27ac, the majority of significant interactions (80% for Naïve HaCaT) do not overlap H3K27ac peaks. This is in stark contrast with results from HiChIP for which, as expected, 99.8% of interactions overlap a H3K27ac peak at one end while 92.5% of interactions overlap a peak at both ends.

Moreover, Capture Hi-C significant interactions are not very specific for active genes: 36% of the genes interacting with baits in naïve HaCaT are not expressed. This is in contrast with HiChIP, for which the majority of interactions are coming from active H3K27ac regions and 82% interacting genes from the same capture Hi-C regions were found to be expressed. This supports that HiChIP results are greatly enriched for active and informative interactions compared to Capture Hi-C. In addition, HiChIP finds significantly more interacting genes than Capture Hi-C (figure 2B). We note that whilst the Capture Hi-C findings and HiChIP findings are complementary, they enrich for different features and therefore represent different viewpoints, with HiChIP giving greater relevance to disease genetics in this particular cell type. In addition, HiChIP allows for interpretation across different immune/skin-mediated diseases whereas Capture Hi-C is mostly disease-specific. Therefore, we focus on HiChIP-implicated genes for the remainder of this study.

Chromatin contacts identified by HiChIP are confirmed by functional evidence

We first wanted to confirm that HiChIP can successfully discovery important regulatory features involved in gene regulation. To do this we interrogated our HaCaT stimulation dataset and used RNA-seq to discover differentially expressed genes. We identify 535 differentially genes which strongly enriched for interferon related pathways (figure S1). We assume that these genes are regulated through enhancers binding transcription factors related to response to the stimulation. As expected, H3K27ac peaks linked through HiChIP significant interactions to differentially expressed genes are enriched for TF motifs that are known to be involved in the response to interferon gamma stimulation (figure 3A).

Next, we wanted to see how HiChIP datasets can recapitulate results from very large eQTL studies in the discovery of genes dysregulated by variants associated with disease. We used psoriasis loci as a proof of concept and identified the genes for which expression levels were influenced by the GWAS SNPs. Using the largest blood eQTL database available, eQTLgen (Võsa *et al.* 2018), we evaluated our results from naïve T cells and GM12878 cells. Our HiChIP analysis showed 51% recall rate and 32% specificity compared to the genes identified from eQTLs. This shows very high concordance despite the fact that blood contains primarily erythrocytes while lymphocytes are a minority.

We then compared our keratinocytes datasets with the GTEX dataset from sun exposed skin (GTEx Consortium 2013). Here the GTEX dataset contains relatively few statistically significant eQTLs compared to the eQTLgen database. Using our data, we see a recall rate of 53% and a specificity of 7.5%. Again, skin also contains more cell lineages than keratinocytes, and the low specificity is due to the lower power of this dataset compared to eQTLgen. Interestingly, 38 out of 52 skin specific eQTLs linked to psoriasis loci are recapitulated in the eQTLgen dataset.

These results, combined, demonstrate that our relatively small HiChIP dataset can successfully recapitulate an eQTL dataset produced with 31000 samples and that we can discover significantly more genes compared to the GTEX database with more than 800 samples, giving us power to identify disease specific interactions in disease relevant cell types.

GWAS variants are significantly enriched in the cell types studied

We wanted to test the assumption that GWAS results from various diseases would enrich for regulatory elements that are active in cell types relevant to their disease mechanism. To do so we used the H3K27ac signal collected from our HiChIP data using our recently developed tool, HiChIP-Peaks (Shi, Rattray and Orozco 2019), to estimate an enrichment score of these SNPs over a background (figure 3B). As expected, SNPs associated with psoriasis, PsA, atopic dermatitis, melanoma and SSc show significant enrichment for the activity marker H3K27ac in the cell lines studied. As a control we used RA, which is characterized by joint inflammation like PsA, but there is an absence of skin involvement. As expected, RA has much stronger enrichment in T cells and B cells relative to keratinocytes compared to the studied dermatological traits, such as psoriasis, psoriatic arthritis and melanoma. Interestingly psoriatic arthritis seems to have a much higher enrichment in keratinocytes than psoriasis even though most of the psoriatic arthritis loci are also shared by psoriasis.

The importance of studying disease specific cell types is evidenced by the fact that many SNPs overlap active regions of the genome only in specific cell types. Taking psoriasis as an example, we see that of the 45 loci that overlap a H3K27ac peak in any cell line, only 22 do so in all cell types, while 4 loci overlap a peak only in keratinocytes. Finally, for all traits studied, INF-γ stimulated cells provided an increased enrichment of H3K27ac levels on the GWAS loci compared to their respective background. In psoriasis, 47 (15%) more SNPs are overlapping H3K27ac peaks in stimulated HaCaT compared to unstimulated providing potential involvement in a significant larger number of SNPs.

Linking genes to disease associated loci provide functional dissection of disease mechanisms

Using the combined HiChIP datasets and cell-type matched RNA-seq data we identified transcribed genes whose TSS either overlap or are linked by chromatin interaction to GWAS SNPs associated with psoriatic arthritis, psoriasis, atopic dermatitis, melanoma and systemic sclerosis. These traits were selected because they have significant contributions derived from keratinocytes and skin infiltrating T cells. As they also have significant involvement from the immune system, we included Naïve T cells and B cells in the analysis.

We identify between 77 and 481 potentially linked genes to each disease studied (table 2). Importantly, these genes were strongly enriched for disease relevant pathways. For example, genes linked to melanoma loci were enriched for replicative senescence and cell cycle pathways, while genes linked to psoriasis and psoriatic arthritis were linked to cytokine signalling and interleukin-23 (figure 4), which is a pathway targeted by multiple novel treatments (Sakkas, Zafiriou and Bogdanos 2019).

We found some differences in disease pathways that indicated the underlying pathology, for example, psoriasis pathways included regulation of cell adhesion mediated by integrin (important to the epithelium), that was not found in PsA: a related disorder that shares much disease background with Ps but primarily affects joints and not always skin.

Although some genes were common in all cell lines (e.g. in psoriasis 18.7%) most genes were specific to one or a few cell types (figure S2). Moreover, we found that most loci also implicate more than one gene, with on average 6.8 genes per loci in psoriasis.

Using this data, we can provide functional insight into the mechanism involved in mediating disease susceptibility. For example, psoriasis and atopic dermatitis have distinct associations at the *ETS1* locus. However, whilst the psoriasis locus directly overlaps the promoter of *ETS1*, the association

with atopic dermatitis is located 130kb downstream of the gene (200kb from the promoter). We observe significant chromatin interactions between the atopic dermatitis locus and the *ETS1* promoter in Naïve T cells, Myla cells and GM12878 cells (figure 5). Our data suggests a putative mechanism for how the distinct disease associations at this locus are mediated by a single gene.

HiChIP identifies novel genes associated with psoriasis loci in a cell-type specific manner

The results from our analysis can be applied to augment our understanding of disease associated loci and link novel genes or change the ones currently linked. For the purpose of this work we focus on 4 psoriasis associated loci which show cell type specific and previously unknown interactions.

The psoriasis locus linked by the SNP rs73178598 is located in an intergenic region overlapping an antisense RNA, *SATB1-AS1*. We found a 240kb T cell specific interaction present in Naïve T cells and MyLa cells but not keratinocytes linking this locus with the promoter of *SATB1* (figure 6). This gene has not been previously linked to psoriasis genetically. Silencing of *SATB1* has been shown to have a similar effect to IFN-γ stimulation on MHC chromatin organization (Pavan Kumar *et al.* 2007) and is known to be an important regulator of Tregs and autoimmunity (Beyer *et al.* 2011).

The psoriasis loci linked by the SNP rs9504361 is intronic to *EXOC2* and is typically associated with this gene. However, despite *EXOC2* being widely expressed and its expression being associated with this SNP in blood, H3K27ac occupancy is specific to MyLa in our analysis. A long-range interaction was detected between this SNP and the promoters of *IRF4* and *DUSP22*, specifically in this cell type and not in Naïve $CD4^+$ T cells or keratinocytes (figure 7). *IRF4* is a lymphocyte specific transcription factor that negatively regulates Toll-like-receptor (TLR) signalling, as part of the interferon response, central to the activation of innate and adaptive immune systems (Huber and Lohoff 2014). It was found to be overexpressed in psoriatic skin lesions (Ni *et al.* 2012). DUSP22 is a phosphatase that might be involved in the JNK signalling pathway and has been shown to be associated with lymphomas (Zeke *et al.* 2016; Paydas *et al.* 2019).

Another example of a novel gene target can be found at the 1p36 locus (rs10794648), which to-date has been associated with *IFNLR1* (also known as IL28-RA) as it is the closest gene to the associated SNPs (Strange *et al.* 2010; Stuart *et al.* 2015). *IFNLR1* protein encodes part of a receptor for IFN-γ that is presented in the epidermis, and thought to promote antiviral response in psoriasis (Lazear, Nice and Diamond 2015). However, our HiChIP data showed long-range interactions between the psoriasis SNPs at 1p36 and the distal gene *GRHL3* specifically in keratinocytes (figure 8). This gene is a transcription factor, upregulated in psoriatic lesions and required for repair of the epidermal skin barrier following immune-mediated injury (Gordon *et al.* 2014). Targets of the *GRHL3* TF include further GWAS-implicated genes *IVL* (pro-differentiation) and *KLF4* (TF; KC differentiation, skin barrier formation). *KLF4* was shown previously as upregulated by IFN- γ and implicated as a likely functional target gene in the 9q31 locus through chromatin looping and CRISPR activation (Ray-Jones *et al.* 2019).

Finally, the locus marked by the psoriasis SNPs rs73183592 with previously unknown function and positioned within an exon of a lncRNA was linked via a long-range interaction spanning about 500kb to *FOXO1* (figure 9), a gene with important functions in regulatory T cells. Interestingly, this interaction was identified only in keratinocytes in our analysis. Dysregulation of this pathway was also found the be important in the development of psoriasis (Li *et al.* 2019; Zhang and Zhang 2019).

Discussion

Chromatin conformation and functional genomics studies have the potential to uncover the underlying mechanisms that drive the disease susceptibility of many complex traits. Although these techniques are very promising, there has been a lack of studies in disease relevant cell types, and as recently evidenced, both chromatin interactions and gene regulation are cell type and stimulation specific (Dixon *et al.* 2012; Rao *et al.* 2014; Schmitt *et al.* 2016; Burren *et al.* 2017; Mumbach *et al.* 2017; Rubin *et al.* 2017; Siersbæk *et al.* 2017; Hansen *et al.* 2018).

Here we used H3K27ac HiChIP, a novel technique that allows combined analysis of both chromatin conformation and chromatin activity, to create the first global study of promoter-enhancer interactions in keratinocytes and tissue resident $CD8^+$ T cells. Equivalent data has so far only been generated in immune cells with few examples in other cell populations, such as HiChIP in endometrial cancer cells (O'Mara, Spurdle and Glubb 2019), and promoter-capture Hi-C in neuronal cells (Song *et al.* 2019), cardiomyocytes (Choy *et al.* 2018) and pancreatic islets (Miguel-Escalada *et al.* 2019). A common limitation of these studies is that they make use of immortalized cell lines or differentiated pluripotent stem cells. Ideally, future studies would involve primary cells or tissue from patients.

We have analysed the effectiveness of these techniques to link functional relevant elements and shown how we can use them to study chromatin interaction and activity in a cell type specific manner. With the combination of public data for cell types relevant in the disorders studied we have analysed the potential information that can be gained for a number of skin-related disorders.

Using these datasets, we study all disease associated SNPs for psoriatic arthritis, psoriasis, atopic dermatitis, melanoma and systemic sclerosis, and identify all the genes that are linked by chromatin interactions to these variants. We show that these genes enrich for disease relevant pathways and provide tables for all loci (supplementary tables). We demonstrate how this data allows us to identify novel mechanisms and we show, using 4 distinct psoriasis associated loci, how we can provide functional insight and link novel genes or change the ones currently linked.

We highlight two loci in which keratinocytes provide us cell type specific interactions that were not found in the more commonly used immune cell populations. We link the locus marked by rs73183592 to *FOXO1*, a gene located about 500kb away from the locus. We also show a link between the psoriasis SNPs rs10794648 to *GRHL3,* a gene involved in the epidermis and fundamental to keratinocytes' function. Phenotypically, psoriasis, is a disease primarily affecting skin barrier and this locus possibly explain some of the genetic background that leads to developing the disease. These novel connections have the potential to be used as therapeutic targets in drug repurposing and discovery, as recently applied for other diseases (Fang *et al.* 2019; Martin *et al.* 2019).

Whilst these results are important, a limitation of our study is that further biological and functional validation is required, such as ones provided by our recent study in the KLF4 locus (Ray-Jones *et al.* 2019). Nevertheless, through our analysis we present a list of potential target genes and pathways for mediating disease risk for 5 complex diseases and by documenting individual loci we highlight the mechanisms by which this risk is mediated. The genes and mechanisms represent a useful resource for further research aimed at characterising how increased disease susceptibility is mediated for these and other complex diseases.

Methods

Cell culture

HaCaT keratinocyte cells were obtained from Addexbio (T0020001); these are in vitro spontaneously transformed keratinocytes from histologically normal skin. Cells were cultured in high-glucose Dulbecco's modified eagle's medium (DMEM) supplemented with 10% foetal bovine serum (FBS) and penicillin-streptomycin (Thermo Fisher Scientific, final concentration 100 U penicillin, 0.1 mg streptomycin/ml). For HaCaT stimulation experiments, the media was supplemented with 100 ng/mL recombinant human IFN-γ (285-IF-100; R&D Systems) and cells incubated for 8 hours prior to harvest.

My-La CD8+ cells were obtained from Sigma-Aldrich (95051033). These cells are cancerous human T-lymphocytes derived from a patient with mycosis fungoides. Cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 10% AB human serum (Sigma Aldrich), 100 U/mL recombinant human IL-2 (Sigma-Aldrich) and penicillin-streptomycin (final concentration 100 U penicillin, 0.1 mg streptomycin/ml).

Cell crosslinking for chromatin-based experiments

HaCaT and My-La cells were crosslinked for 10 minutes in 1% (HiChIP) or 2% (Hi-C) formaldehyde. The reaction was quenched with 0.135M glycine, the supernatant removed, and the cells snap frozen on dry ice and stored at -80°C.

HiChIP library generation and processing

HiChIP libraries were generated according to the Chang Lab protocol (Mumbach *et al.* 2016). Briefly, 10 million crosslinked cells were lysed, digested with MboI, biotinylated and ligated. Immunoprecipitation was performed with H3K27ac antibody, after which the DNA was purified and de-crosslinked followed by biotin-streptavidin pulldown. Tagmentation was performed using Illumina Tn5 and the libraries were amplified with Nextera indexing primers. Sequencing was performed using paired-end Illumina SBS technology.

Sequencing data for the HiChIP libraries was filtered and the adapters were removed using fastp v0.19.4 (Chen *et al.* 2018). The reads were then mapped to the GRCh38 genome with Hi-C Pro v2.11.0 (Servant *et al.* 2015), using default settings. Enriched regions (H3K27ac peaks) where identified using HiChIP-peaks v 0.1.1 (Shi, Rattray and Orozco 2019) with default settings and FDR<0.01. Loops were identified using FitHiChIP (Bhattacharyya *et al.* 2019) using the following settings: Coverage normalization, stringent background with merging enabled, peaks generated from HiChIP-peaks and 5kb bin size.

HiChIP clustering and principal component analysis

To validate the reproducibility and cell type specificity of our HiChIP loops we collected the top 10000 significant loops from each individual replicate and combined it to create a set of 82545 unique loops across all samples. For each of these loops we then collected the raw FitHiChIP p-value from each sample from the raw interactions file (replacing missing entries with 1). We then run hierarchical clustering using correlation (seaborn clustermap) and PCA (Scikit-learn) analysis on the resulting data matrix.

For the clustering of the peaks we used the included differential peak calling module of HiChIP-peaks (Shi, Rattray and Orozco 2019) and used the data matrix provided to run hierarchical clustering using Euclidean distance.

Hi-C library generation and processing

In-situ Hi-C libraries were generated as previously described as a pre-Capture Hi-C library. Briefly, 50 million crosslinked cells were lysed and the chromatin digested with HindIII, followed by biotinylation and in-nucleus ligation. Crosslinks were reversed, the DNA was purified and biotin-streptavidin pulldown was used to enrich for ligation sites. Hi-C libraries were sequenced using paired-end Illumina SBS technology (see Supplementary Methods for details). The Sequencing data was filtered and adapters were removed using fastp v0.19.4 (Chen *et al.* 2018). The reads were then mapped to the GRCh38 genome with Hi-C Pro v2.11.0 (Servant *et al.* 2015), using default settings. The Hi-C interaction matrices were normalised within Hi-C Pro using iterative correction and eigenvector decomposition (ICE). TADs were identified using OnTAD v1.2 (An *et al.* 2019), a novel Optimized Nested TAD caller for Hi-C data, using Hi-C data binned at a 40kb resolution and a maximum TAD size of 4mb.

Region Capture Hi-C and overlap with HiChIP results

Region capture Hi-C libraries were generated as previously described from the Hi-C libraries as part of a previous study (Ray-Jones *et al.* 2019). Briefly, RNA baits were designed to target all known non-MHC GWAS loci for a number of diseases including psoriasis, juvenile idiopathic arthritis, asthma, psoriatic arthritis, rheumatoid arthritis and systemic sclerosis. Each 120 bp bait was targeted to within 400 bp of a HindIII fragment end, comprised 25-65% GC content and contained fewer than three unknown bases. The baits were synthesised by Agilent Technologies.

For each cell type 50 million crosslinked cells were lysed and the chromatin digested with HindIII, followed by biotinylation and in-nucleus ligation. Crosslinks were reversed, the DNA was purified and biotin-streptavidin pulldown was used to enrich for ligation sites. The libraries were amplified and capture was performed using the RNA baits described above using the SureSelect reagents and protocol. Following a second amplification the libraries were sequenced using paired-end Illumina SBS technology.

Capture Hi-C sequence data was quality filtered with fastp v 0.19.4 (Chen *et al.* 2018) and then processed through the Hi-C User Pipeline (HiCUP) v0.7.2 (Wingett *et al.* 2015) and mapped to the GRCh38 genome. For each cell type, the two biological replicates were simultaneously run through Capture Hi-C Analysis of Genomic Organisation (CHiCAGO) v1.10.1 (Cairns *et al.* 2016) in R v3.5.1 and significant interactions were called with a score threshold of 5.

To identify active enhancer-promoter interactions we kept the CHiCAGO interactions that originated from HiChIP H3K27ac peaks in the matching cell type. We then identified the expressed promoters (TPM>1) that were within 5kb of the other end of the interactions.

To compare the results from Capture Hi-C with our new HiChIP libraries we determined the interactions that originated from within 5kb of those 4500 capture fragments. Genes were then identified as described in “Linking GWAS results to putative gene targets” section. The resulting genes were then compared with the results from Capture Hi-C.

RNA-seq

3’ mRNA sequencing libraries were generated for cell lines using the Lexogen QuantSeq 3’ mRNA-Seq Library Prep Kit FWD for Illumina. Libraries were sequenced using single-end Illumina SBS technology. Reads were quality trimmed using Trimmomatic v0.38 (Bolger, Lohse and Usadel 2014) using a sliding window of 5 with a mean minimum quality of 20. Adapters and poly A/poly G tails were removed using Cutadapt v1.18 (Martin 2011) and then UMIs were extracted from the 5’ of the reads using UMI-tools v0.5.5 (Smith, Heger and Sudbery 2017). Reads were then mapped using STAR v2.5.3a (Dobin *et al.* 2013) on the GRCh38 genome with GENCODE annotation v29 (Harrow *et al.* 2012). Reads were de-

duplicated using UMIs with UMI-tools and then counted using HTSeq v0.11.2 (Anders, Pyl and Huber 2015). Count matrixes were analysed in R 3.5.1 and normalisation and differential expression analysis was conducted using DESeq2 v1.22.2 (Love, Huber and Anders 2014). Differentially expressed genes were called with an adjusted P value of 0.10 (FDR 10%). For detection of expressed genes in the cell lines, we considered RNA-seq counts greater than 1 count per million.

Public RNA-seq data

Public RNA-seq for the CD4 naïve t cell type was downloaded from (Bonnal *et al.* 2015). Accession ID: ERP004883

https://www.ncbi.nlm.nih.gov/pubmed/26451251?dopt=Abstract&holding=npg

Raw sequencing reads were filtered and adapters and polyAs trimmed with fastp v 0.19.4 (Chen *et al.* 2018). Reads were then mapped with salmon v0.14.1 (Patro *et al.* 2017) to the GRCh38 genome with GENCODE annotation v29 (Harrow *et al.* 2012).

TPM values were used later in the analysis for gene filtering.

GWAS data

Genome wide significant (p-value < 5x10^-8) GWAS loci were downloaded for the following diseases: psoriatic arthritis (Bowes *et al.* 2015; Stuart *et al.* 2015), psoriasis (Tsoi *et al.* 2017), melanoma (Duffy *et al.* 2018), systemic sclerosis (López-Isac *et al.* 2019), atopic dermatitis (Paternoster *et al.* 2015) and rheumatoid arthritis (Okada *et al.* 2014).

SNPs in high linkage disequilibrium (R^2 > 0.8) with the lead SNPs were identified using plink v1.90b3.39 on the 1000 genomes data v3 with population set to EUR.

SNP enrichment

We obtained the H3K27ac signal tracks for each cell type from the HiChIP data using HiChIP-Peaks. This track corresponds to the signal for this marker of activity along the genome. We then calculated the median intensity the signal over every SNP outside of the MHC and compared it with the median for a set of 1 million randomly generated positions to get an estimate of a genomic background for the signal to calculate an enrichment. We also calculated the number of individual SNPs that are located within a H3K27ac peak for each cell type.

Linking stimulation responsive genes to enhancers

We identified the genes that were differentially expressed during the INFγ stimulation as described in the other section. We then identified the promoter regions for these genes and identified the genomic regions that interacted with these promoters in any of the two conditions in HaCaT cells.

We intersected these regions with the H3K27ac peaks identified by HiChIP-Peaks to narrow down the exact location of the interacting enhancers and we run Motif enrichment analysis using HOMER v 4.8.3 (Heinz *et al.* 2010) with the findMotifsGenome.pl command and "–size given" parameter.

Linking GWAS loci to putative gene targets

To identify the genes that were linked to disease associated SNPs we first identified the transcription start sites of all protein coding transcripts in the hg38 Gencode V29 annotation (Harrow *et al.* 2012). We associated all transcripts for which promoter was located within 5kb of a loop. We also associated the transcripts for which the TSS was within 1kb of a SNP overlapping a H3K27ac peak as identified from HiChIP data.

All transcripts were then grouped by originating gene and all analysis was done at the gene level.

Lastly genes were filtered by expression level of at least 1 TPM in the corresponding cell line.

Overlap with eQTL

We downloaded the full cis-eQTL datasets from the sun-exposed skin GTEx v7 dataset (GTEx Consortium 2013) and the eQTLgen (version 2018/10/17) dataset (Võsa *et al.* 2018).

To identify the eQTLs that were originating from the GWAS loci we simply queried every SNP that was in LD with the lead SNP and recorded all the genes that were significantly linked to those variants. Genes were filtered by expression TPM > 1 in our cell types.

We then identified all the genes that are linked for a specific disease (in this example psoriasis) and compared this list with the list of genes that were identified from the same SNPs using the HiChIP interactions.

Pathway analysis

The most enriched pathways for each disease were identified using the EnrichR (Chen *et al.* 2013) web API with the gene set library set to GO_Biological_Process_2018.

The pathways were then sorted by p-value and the top 10 enriched pathways were plotted.

Acknowledgements

The authors would like to acknowledge the assistance given by IT Services and the use of the Computational Shared Facility at The University of Manchester.

This work was funded by the Wellcome Trust (award references 207491/Z/17/Z and 215207/Z/19/Z), Versus Arthritis (award reference 21754), NIHR Manchester BRC and the Medical Research Council (award reference MR/N00017X/1).

Conflict of Interest

None declared.

References

Aguet F, Brown AA, Castel SE *et al.* Genetic effects on gene expression across human tissues. *Nature* 2017;**550**:204–13.

Albanesi C, Madonna S, Gisondi P *et al.* The interplay between keratinocytes and immune cells in the pathogenesis of psoriasis. *Front Immunol* 2018;**9**, DOI: 10.3389/fimmu.2018.01549.

An L, Yang T, Yang J *et al.* Hierarchical Domain Structure Reveals the Divergence of Activity among TADs and Boundaries. *bioRxiv* 2019:361147.

Anders S, Pyl PT, Huber W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015;**31**:166–9.

Antohe M, Nedelcu RI, Nichita L *et al.* Tumor infiltrating lymphocytes: The regulator of melanoma evolution (Review). *Oncol Lett* 2019;**17**:4155–61.

Beyer M, Thabet Y, Müller RU *et al.* Repression of the genome organizer SATB1 in regulatory T cells is required for suppressive function and inhibition of effector differentiation. *Nat Immunol* 2011;**12**:898–907.

Bhattacharyya S, Chandra V, Vijayanand P *et al.* Identification of significant chromatin contacts from HiChIP data by FitHiChIP. *Nat Commun* 2019;**10**, DOI: 10.1038/s41467-019-11950-y.

Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;**30**:2114–20.

Bonnal RJP, Ranzani V, Arrigoni A *et al.* De novo transcriptome profiling of highly purified human lymphocytes primary cells. *Sci Data* 2015;**2**:150051.

Bowes J, Budu-Aggrey A, Huffmeier U *et al.* Dense genotyping of immune-related susceptibility loci reveals new insights into the genetics of psoriatic arthritis. *Nat Commun* 2015;**6**:6046.

Burren OS, Rubio García A, Javierre B-M *et al.* Chromosome contacts in activated T cells identify autoimmune disease candidate genes. *Genome Biol* 2017;**18**:165.

Cai Y, Fleming C, Yan J. New insights of T cells in the pathogenesis of psoriasis. *Cell Mol Immunol* 2012;**9**:302–9.

Cairns J, Freire-Pritchett P, Wingett SW *et al.* CHiCAGO: robust detection of DNA looping interactions in Capture Hi-C data. *Genome Biol* 2016;**17**:127.

Chen EY, Tan CM, Kou Y *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 2013;**14**:128.

Chen S, Zhou Y, Chen Y *et al.* fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;**34**:i884–90.

Choy M-K, Javierre BM, Williams SG *et al.* Promoter interactome of human embryonic stem cell-derived cardiomyocytes connects GWAS regions to cardiac gene networks. *Nat Commun* 2018;**9**:2526.

Dixon JR, Selvaraj S, Yue F *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 2012;**485**:376–80.

Dobin A, Davis CA, Schlesinger F *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;**29**:15–21.

Dryden NH, Broome LR, Dudbridge F *et al.* Unbiased analysis of potential targets of breast cancer

susceptibility loci by Capture Hi-C. *Genome Res* 2014;**24**:1854–68.

Duffy DL, Zhu G, Li X *et al.* Novel pleiotropic risk loci for melanoma and nevus density implicate multiple biological pathways. *Nat Commun* 2018;**9**, DOI: 10.1038/s41467-018-06649-5.

Ernst J, Kheradpour P, Mikkelsen TS *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 2011;**473**:43–9.

Fang H, Beckmann G, Bountra C *et al.* A genetics-led approach defines the drug target landscape of 30 immune-related traits. *Nat Genet* 2019, DOI: 10.1038/s41588-019-0456-1.

Farh KK-H, Marson A, Zhu J *et al.* Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 2015;**518**:337–43.

Gordon WM, Zeller MD, Klein RH *et al.* A GRHL3-regulated repair pathway suppresses immune-mediated epidermal hyperplasia. *J Clin Invest* 2014;**124**:5205–18.

GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;**45**:580–5.

Hansen AS, Cattoglio C, Darzacq X *et al.* Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus* 2018;**9**:20–32.

Harrow J, Frankish A, Gonzalez JM *et al.* GENCODE: The reference human genome annotation for the ENCODE project. *Genome Res* 2012;**22**:1760–74.

Heinz S, Benner C, Spann N *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell* 2010;**38**:576–89.

Hennino A, Vocanson M, Toussaint Y *et al.* Skin-Infiltrating CD8 + T Cells Initiate Atopic Dermatitis Lesions . *J Immunol* 2007;**178**:5571–7.

Hijnen D, Knol EF, Gent YY *et al.* CD8+ T cells in the lesional skin of atopic dermatitis and psoriasis patients are an important source of IFN-γ, IL-13, IL-17, and IL-22. *J Invest Dermatol* 2013;**133**:973–9.

Huber M, Lohoff M. IRF4 at the crossroads of effector T-cell fate decision. *Eur J Immunol* 2014;**44**:1886–95.

Jäger R, Migliorini G, Henrion M *et al.* Capture Hi-C identifies the chromatin interactome of colorectal cancer risk loci. *Nat Commun* 2015;**6**:6178.

Javierre BM, Burren OS, Wilder SP *et al.* Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. *Cell* 2016;**167**:1369-1384.e19.

Kundaje A, Meuleman W, Ernst J *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* 2015;**518**:317–30.

Lazear HM, Nice TJ, Diamond MS. Interferon-λ: Immune Functions at Barrier Surfaces and Beyond. *Immunity* 2015;**43**:15–28.

Lees JR. Interferon gamma in autoimmunity: A complicated player on a complex stage. *Cytokine* 2015;**74**:18–26.

Li B, Lei J, Yang L *et al.* Dysregulation of Akt-FOXO1 Pathway Leads to Dysfunction of Regulatory T Cells in Patients with Psoriasis. *J Invest Dermatol* 2019;**139**:2098–107.

López-Isac E, Acosta-Herrera M, Kerick M *et al.* GWAS for systemic sclerosis identifies multiple risk loci and highlights fibrotic and vasculopathy pathways. *Nat Commun* 2019;**10**, DOI:

10.1038/s41467-019-12760-y.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;**15**:550.

Lowes MA, Kikuchi T, Fuentes-Duculan J *et al.* Psoriasis vulgaris lesions contain discrete populations of Th1 and Th17 T cells. *J Invest Dermatol* 2008;**128**:1207–11.

Mahajan A, Taliun D, Thurner M *et al.* Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet* 2018;**50**:1505–13.

Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011;**17**:10.

Martin P, Ding J, Duffus K *et al.* Chromatin interactions reveal novel gene targets for drug repositioning in rheumatic diseases. *Ann Rheum Dis* 2019;**78**:1127–34.

Martin P, McGovern A, Massey J *et al.* Identifying Causal Genes at the Multiple Sclerosis Associated Region 6q23 Using Capture Hi-C. *PLoS One* 2016;**11**:e0166923.

Martin P, McGovern A, Orozco G *et al.* Capture Hi-C reveals novel candidate genes and complex long-range interactions with related autoimmune risk loci. *Nat Commun* 2015;**6**:1–7.

Mccoy SS, Reed TJ, Berthier CC *et al.* Scleroderma keratinocytes promote fibroblast activation independent of transforming growth factor beta. , DOI: 10.1093/rheumatology/kex280.

McGovern A, Schoenfelder S, Martin P *et al.* Capture Hi-C identifies a novel causal gene, IL20RA, in the pan-autoimmune genetic susceptibility region 6q23. *Genome Biol* 2016;**17**, DOI: 10.1186/s13059-016-1078-x.

Miguel-Escalada I, Bonàs-Guarch S, Cebola I *et al.* Human pancreatic islet three-dimensional chromatin architecture provides insights into the genetics of type 2 diabetes. *Nat Genet* 2019;**51**:1137–48.

Mizoguchi F, Slowikowski K, Wei K *et al.* Functionally distinct disease-associated fibroblast subsets in rheumatoid arthritis. *Nat Commun* 2018;**9**:1–11.

Mumbach MR, Rubin AJ, Flynn RA *et al.* HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods* 2016;**13**:919–22.

Mumbach MR, Satpathy AT, Boyle EA *et al.* Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat Genet* 2017;**49**:1602–12.

Ni A, Chen H, Wu Y *et al.* Expression of IRF-4 and IBP in skin lesions of patients with psoriasis vulgaris. *J Huazhong Univ Sci Technol - Med Sci* 2012;**32**:287–90.

Nica AC, Dermitzakis ET. Expression quantitative trait loci: present and future. *Philos Trans R Soc Lond B Biol Sci* 2013;**368**:20120362.

O'Mara TA, Spurdle AB, Glubb DM. Analysis of Promoter-Associated Chromatin Interactions Reveals Biologically Relevant Candidate Target Genes at Endometrial Cancer Risk Loci. *Cancers (Basel)* 2019;**11**:1440.

O'reilly S, Hü Gle T, Van Laar JM. T cells in systemic sclerosis: a reappraisal. , DOI: 10.1093/rheumatology/kes090.

Okada Y, Wu D, Trynka G *et al.* Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature* 2014;**506**:376–81.

Paternoster L, Standl M, Waage J *et al.* Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nat Genet* 2015;**47**:1449–56.

Patro R, Duggal G, Love MI *et al.* Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* 2017;**14**:417–9.

Pavan Kumar P, Bischof O, Purbey PK *et al.* Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. *Nat Cell Biol* 2007;**9**:45–56.

Paydas S, Bagir EK, Ergin M *et al.* THE ROLE OF DUSP22 (DUAL SPECIFICITY PHOSPHATASE 22) GENE EXPRESSION IN THE PROGNOSIS OF LOW GRADE LYMPHOMAS. *Hematol Oncol* 2019;**37**:340–1.

Rao SSP, Huntley MH, Durand NC *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 2014;**159**:1665–80.

Ray-Jones H, Duffus K, McGovern A *et al.* Chromatin-based techniques map DNA interaction landscapes in psoriasis susceptibility loci and highlight KLF4 as a target gene in 9q31. *bioRxiv* 2019;**53**:822304.

Rowley MJ, Corces VG. Organizational principles of 3D genome architecture. *Nat Rev Genet* 2018:1.

Rubin AJ, Barajas BC, Furlan-Magaril M *et al.* Lineage-specific dynamic and pre-established enhancer–promoter contacts cooperate in terminal differentiation. *Nat Genet* 2017;**49**:1522–8.

Sakkas LI, Zafiriou E, Bogdanos DP. Mini Review: New Treatments in Psoriatic Arthritis. Focus on the IL-23/17 Axis. *Front Pharmacol* 2019;**10**:872.

Schlaak JF, Buslau M, Jochum W *et al.* T cells involved in psoriasis vulgaris belong to the Th1 subset. *J Invest Dermatol* 1994;**102**:145–9.

Schmitt AD, Hu M, Jung I *et al.* A Compendium of Chromatin Contact Maps Reveals Spatially Active Regions in the Human Genome. *CellReports* 2016;**17**:2042–59.

Servant N, Varoquaux N, Lajoie BR *et al.* HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol* 2015;**16**:259.

Shi C, Rattray M, Orozco G. HiChIP-Peaks: A HiChIP peak calling algorithm. *bioRxiv* 2019:682781.

Siersbæk R, Madsen JGS, Javierre BM *et al.* Dynamic Rewiring of Promoter-Anchored Chromatin Loops during Adipocyte Differentiation. *Mol Cell* 2017;**66**:420-435.e5.

Smith T, Heger A, Sudbery I. UMI-tools: Modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res* 2017;**27**:491–9.

Song M, Yang X, Ren X *et al.* Mapping cis-regulatory chromatin contacts in neural cells links neuropsychiatric disorder risk variants to target genes. *Nat Genet* 2019;**51**:1252–62.

Strange A, Capon F, Spencer CCA *et al.* A genome-wide asociation study identifies new psoriasis susceptibility loci and an interaction betwEn HLA-C and ERAP1. *Nat Genet* 2010;**42**:985–90.

Stuart PE, Nair RP, Tsoi LC *et al.* Genome-wide Association Analysis of Psoriatic Arthritis and Cutaneous Psoriasis Reveals Differences in Their Genetic Architecture. *Am J Hum Genet* 2015;**97**:816–36.

Tsoi LC, Stuart PE, Tian C *et al.* Large scale meta-analysis characterizes genetic architecture for common psoriasis associated variants. *Nat Commun* 2017;**8**:15382.

Uyemura K, Yamamura M, Fivenson DF *et al.* The cytokine network in lesional and lesion-free psoriatic skin is characterized by a t-helper type 1 cell-mediated response. *J Invest Dermatol* 1993;**101**:701–5.

Võsa U, Claringbould A, Westra H-J *et al.* Unraveling the polygenic architecture of complex traits using blood eQTL metaanalysis. *bioRxiv* 2018:447367.

Wingett S, Ewels P, Furlan-Magaril M *et al.* HiCUP: Pipeline for mapping and processing Hi-C data. *F1000Research* 2015;**4**, DOI: 10.12688/f1000research.7334.1.

Wu M, Assassi S. The role of type 1 interferon in systemic sclerosis. *Front Immunol* 2013;**4**, DOI: 10.3389/fimmu.2013.00266.

Zeke A, Misheva M, Reményi A *et al.* JNK Signaling: Regulation and Functions Based on Complex Protein-Protein Partnerships. *Microbiol Mol Biol Rev* 2016;**80**:793–835.

Zhang M, Zhang X. The role of PI3K/AKT/FOXO signaling in psoriasis. *Arch Dermatol Res* 2019;**311**:83–91.

Tables and figures

Table 1. Summary statistics of the datasets used in this project.

Cell Type	Valid interaction pairs (millions)	Median loop distance (kb)	Number of significant loops
HaCaT Naïve (Keratinocytes)	273.1	220	58268
HaCaT INF-γ stimulated	309.8	235	68887
MyLa (CD8+ T cells)	288.6	300	51274
CD4+ Naïve T cells	167.8	310	42710
GM12878 (B-cell)	241.3	185	76451

Table 2. Number of genes identified associated for each disease studied.

Disease	Number of genes identified
PsA	77
Ps	374
Eczema	150
Melanoma	157
Scleroderma	171

Figure 1. A) hierarchical clustering of the loops for the individual HiChIP samples using correlation. B) PCA of the loops for the individual HiChIP samples. C) hierarchical clustering of the HiChIP peaks for the individual samples using Euclidean distance. D) Distance distribution of the significant fithichip interactions for Naïve HaCaT HiChIP.

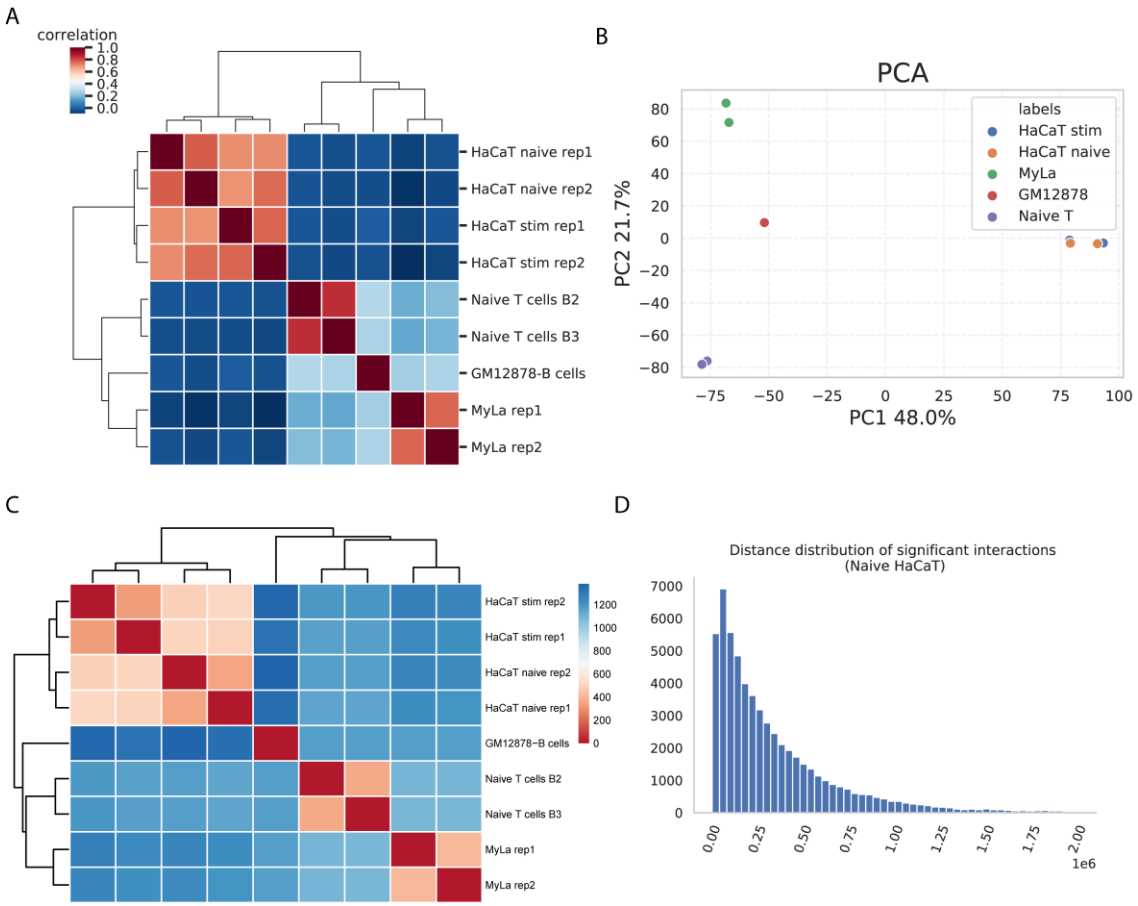


Figure 2. A) Distance distribution of the significant CHiCAGO interactions for Naïve HaCaT region capture Hi-C. B) Overlap of the genes identified through region capture Hi-C and HiChIP in Naïve HaCaT and MyLa cells for the same regions. C) Proportion of significant interactions that are within TADs and that cross TAD boundaries for region capture Hi-C data and HiChIP data.

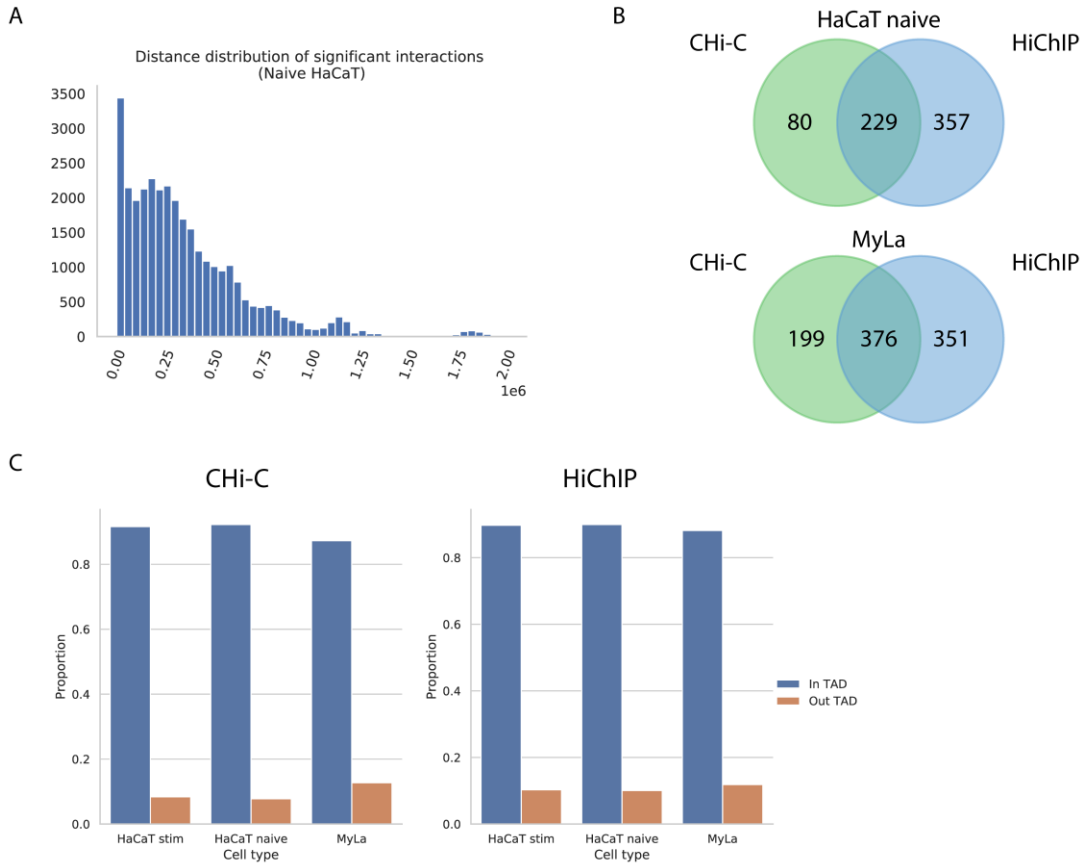


Figure 3. A) Pathways identified from differentially bound HiChIP H3K27ac peaks and HiChIP H3K27ac peaks linked to genes by HiChIP interactions to promoters of genes that were found to be differentially expressed during INFγ stimulation. B) Heatmap showing the enrichment that disease associated variants have in HiChIP H3K27ac signal. Numbers inside each square represent the number of disease associated variants that directly overlap HiChIP H3K27ac peaks in each of the studied cell types. C) Venn diagram showing the number of psoriasis loci which have at least one variant overlapping a HiChIP H3K27ac peak in each cell type.

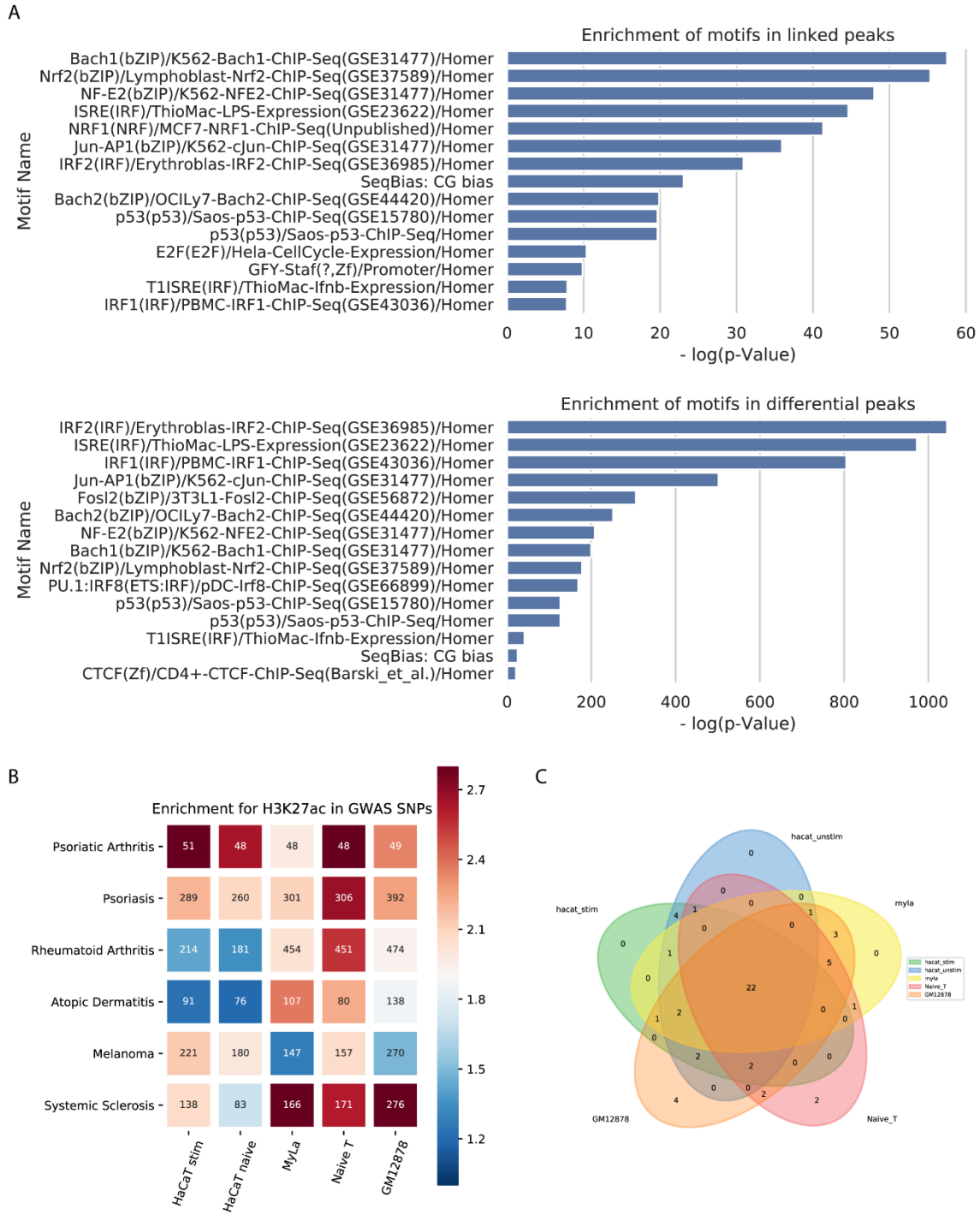


Figure 4. Top pathways enriched by the genes linked to disease associated loci from all cell types.

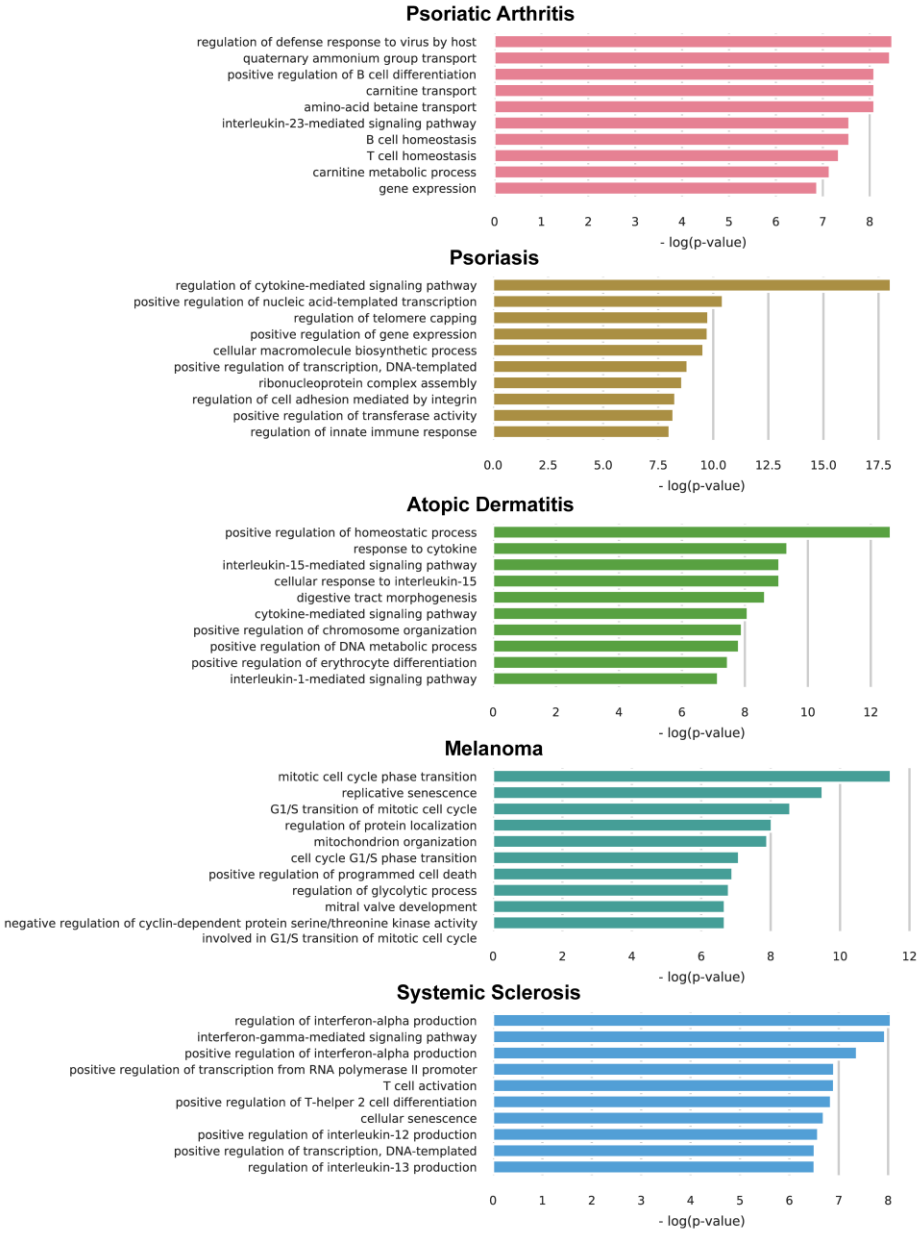


Figure 5. HiChIP interactions from the ETS1 locus provide a functional mechanism for this locus in atopic dermatitis. Tracks (in order): RefSeq genes; SNPs associated with Atopic Dermititis; SNPs associated with Psoriasis; H3K27ac signal from HiChIP-peaks in naïve T cells; Significant long range interactions originating from the atopic dermatitis associated locus from fithichip in naïve T cells; H3K27ac signal from HiChIP-peaks in MyLa cells; Significant long range interactions originating from the atopic dermatitis associated locus from fithichip in MyLa cells; H3K27ac signal from HiChIP-peaks in GM12878 cells; Significant long range interactions originating from the atopic dermatitis associated locus from fithichip in GM12878 cells.

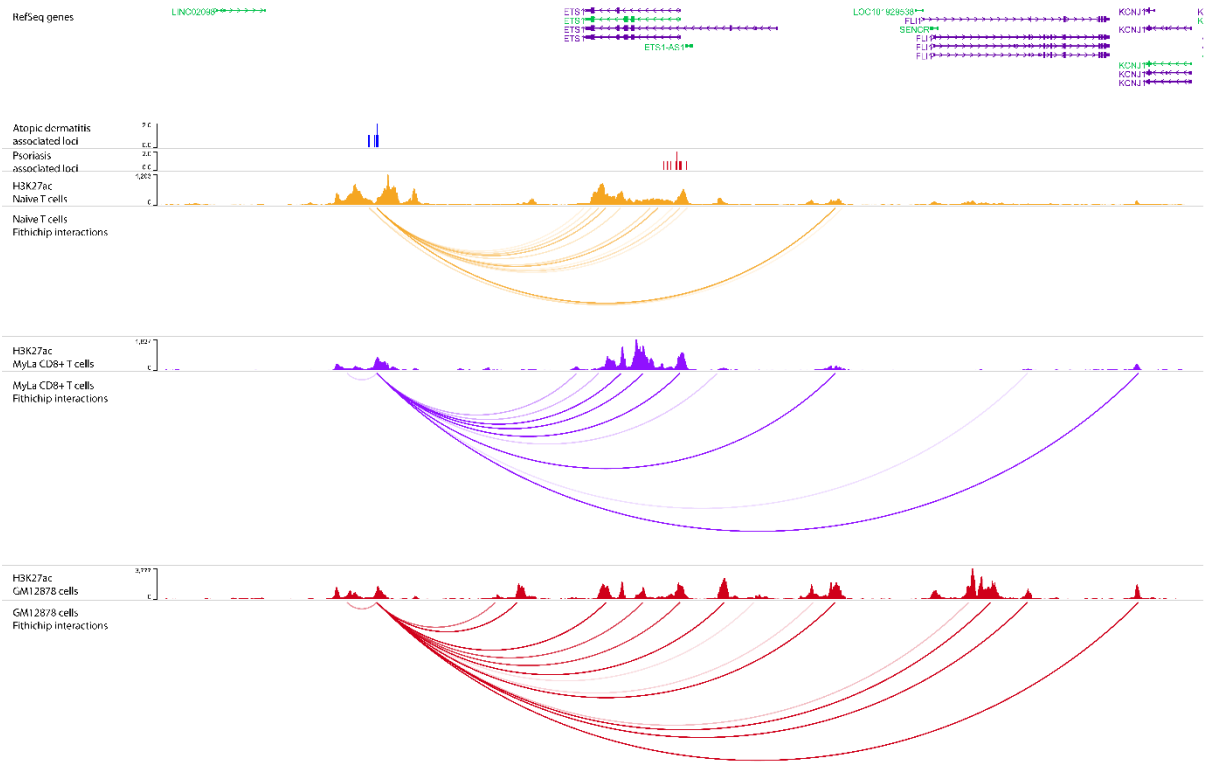


Figure 6. HiChIP interactions from the SATB1/KCNH8 locus link SATB1 as the target gene in this locus. Tracks (in order): RefSeq genes; SNPs associated with Psoriasis; H3K27ac signal from HiChIP-peaks in naïve T cells; Significant long range interactions originating from the psoriasis associated locus from fithichip in naïve T cells; H3K27ac signal from HiChIP-peaks in MyLa cells; Significant long range interactions originating from the psoriasis associated locus from fithichip in MyLa cells.

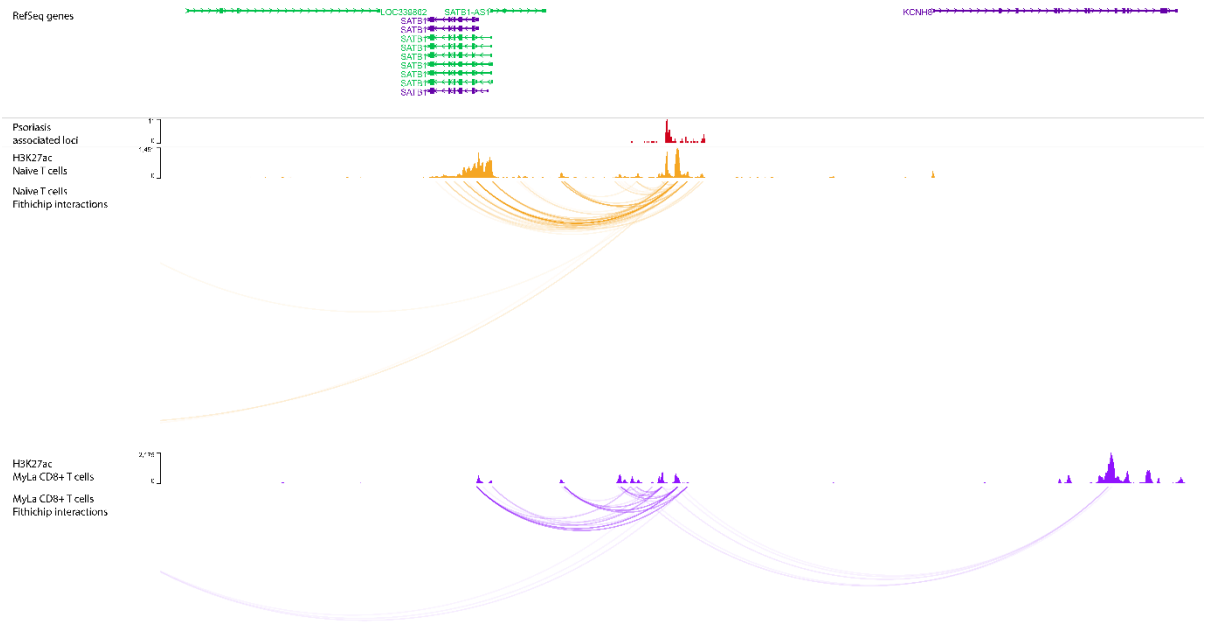


Figure 7. HiChIP interactions from the EXOC2/IRF4/DUSP22 locus link IRF4 and DUSP22 as candidate genes in psoriasis. Tracks (in order): RefSeq genes; SNPs associated with Psoriasis; H3K27ac signal from HiChIP-peaks in naïve T cells; H3K27ac signal from HiChIP-peaks in MyLa cells; Significant long range interactions originating from the atopic dermatitis associated locus from fithichip in MyLa cells.

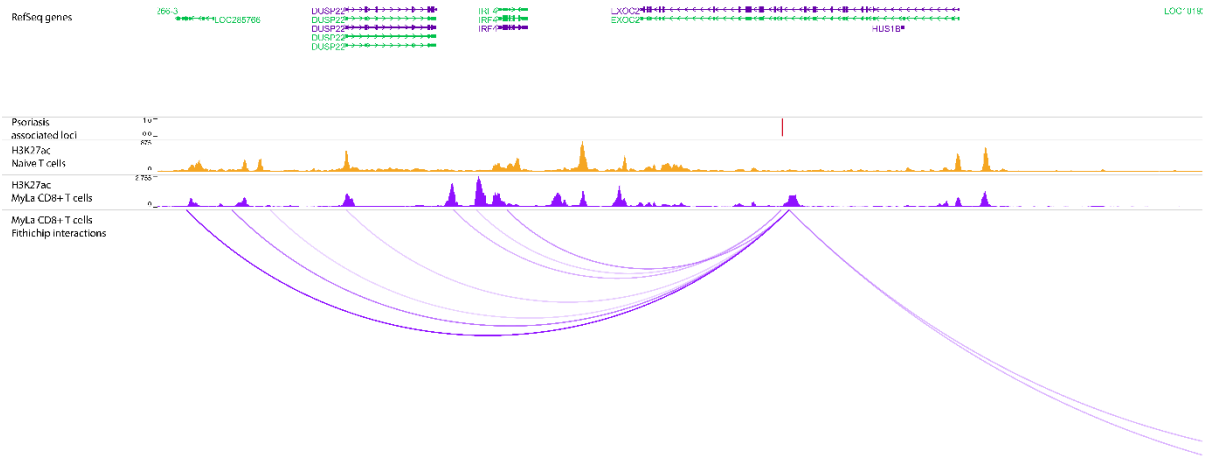


Figure 8. HiChIP interactions from the IFNLR1/GRHL3 locus link GRHL3 as a candidate gene in psoriasis. Tracks (in order): RefSeq genes; SNPs associated with Psoriasis; H3K27ac signal from HiChIP-peaks in naïve HaCaT cells; Significant long range interactions originating from the psoriasis associated locus from fithichip in naïve HaCaT cells; H3K27ac signal from HiChIP-peaks in INFγ stimulated HaCaT cells; Significant long range interactions originating from the psoriasis associated locus from fithichip in INF-γ stimulated HaCaT cells.

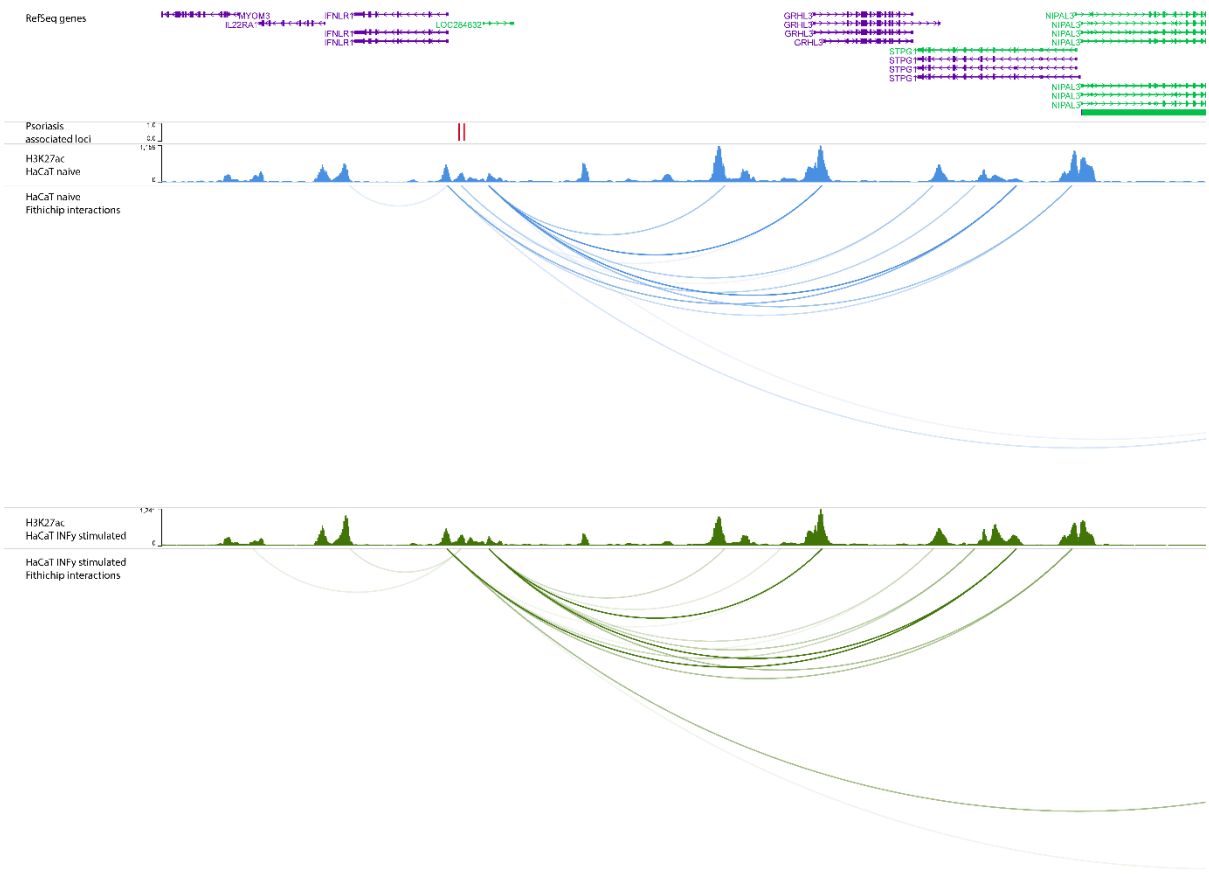
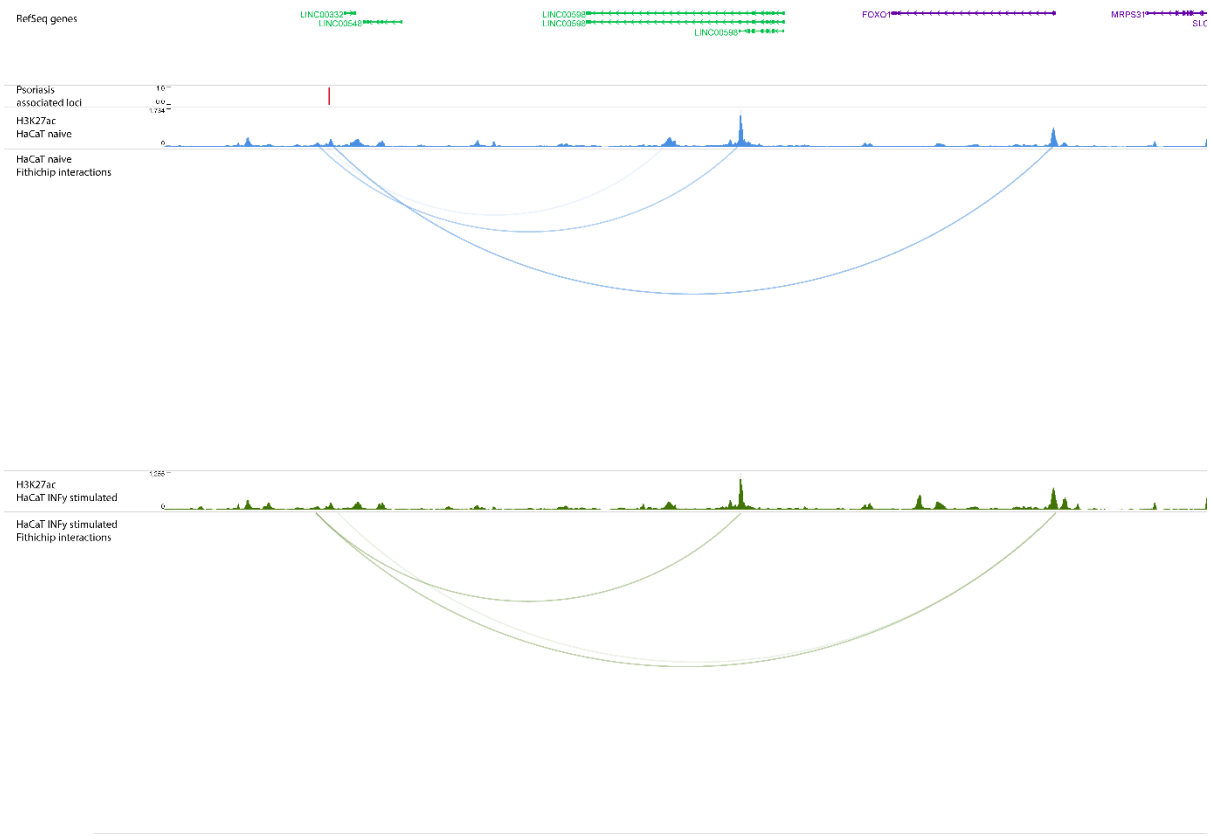


Figure 9. HiChIP interactions from the FOXO1 locus link FOXO1 as a candidate gene in psoriasis. Tracks (in order): RefSeq genes; SNPs associated with Psoriasis; H3K27ac signal from HiChIP-peaks in naïve HaCaT cells; Significant long range interactions originating from the psoriasis associated locus from fithichip in naïve HaCaT cells; H3K27ac signal from HiChIP-peaks in INFγ stimulated HaCaT cells; Significant long range interactions originating from the psoriasis associated locus from fithichip in INF-γ stimulated HaCaT cells.



Supplementary figures for

An active chromatin interactome elucidates the biological mechanisms underlying genetic risk factors of dermatological conditions in disease relevant cell lines

Figure S1. Pathways identified by DE genes it INFgamma stim

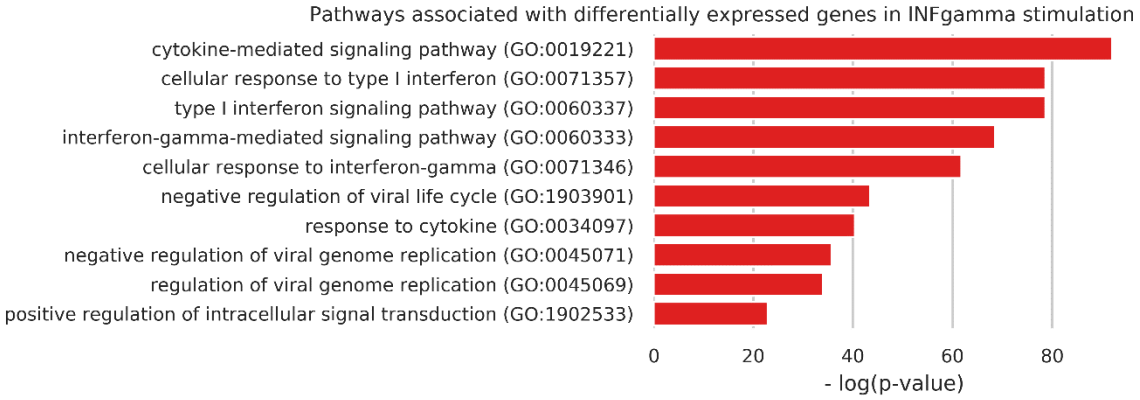


Figure S2. Genes identified for psoriasis by cell type.