

Chromatin Looping Links Target Genes with Genetic Risk Loci for Dermatological Traits



JID Open

Chenfu Shi^{1,8}, Helen Ray-Jones^{1,2,8}, James Ding¹, Kate Duffus¹, Yao Fu³, Vasanthi Priyadarshini Gaddi¹, Oliver Gough¹, Jenny Hankinson⁴, Paul Martin^{1,5}, Amanda McGovern¹, Annie Yarwood^{1,2}, Patrick Gaffney³, Steve Eyre^{1,6}, Magnus Rattray⁷, Richard B. Warren² and Gisela Orozco^{1,6}

Chromatin looping between regulatory elements and gene promoters presents a potential mechanism whereby disease risk variants affect their target genes. In this study, we use H3K27ac HiChIP, a method for assaying the active chromatin interactome in two cell lines: keratinocytes and skin lymphoma-derived CD8+ T cells. We integrate public datasets for a lymphoblastoid cell line and primary CD4+ T cells and identify gene targets at risk loci for skin-related disorders. Interacting genes enrich for pathways of known importance in each trait, such as cytokine response (psoriatic arthritis and psoriasis) and replicative senescence (melanoma). We show examples of how our analysis can inform changes in the current understanding of multiple psoriasis-associated risk loci. For example, the variant rs10794648, which is generally assigned to *IFNLR1*, was linked to *GRHL3*, a gene essential in skin repair and development, in our dataset. Our findings, therefore, indicate a renewed importance of skin-related factors in the risk of disease.

Journal of Investigative Dermatology (2021) **141**, 1975–1984; doi:10.1016/j.jid.2021.01.015

INTRODUCTION

GWASs have uncovered the genetic factors that contribute to disease risk for many complex disorders, including dermatological conditions such as psoriasis (Ps) (Tsoi et al., 2017) or atopic dermatitis (AD) (Paternoster et al., 2015). It is now accepted that the majority of these genetic risk factors do not influence coding sequences directly but rather regulatory regions such as enhancers and promoters, which can be highly cell-type specific (Ernst et al., 2011; Farh et al., 2015;

Roadmap Epigenomics Consortium et al., 2015). Many studies have also shown that the effect of these variants is not necessarily mediated by the closest gene because they can affect distal genes through chromatin looping mechanisms (GTEx Consortium, 2013; GTEx Consortium et al., 2017; Javierre et al., 2016; Nica and Dermitzakis, 2013; Rao et al., 2014; Vösa et al., 2018¹).

Recently, there has been a growing interest in using chromatin conformation and other functional genomics techniques to investigate how these disease-associated loci lead to disease and identify affected genes. To obtain the resolution necessary to identify enhancer–promoter interactions, several techniques have been developed that couple Hi-C methods with various enrichment strategies to increase resolution and reduce sequencing cost compared with traditional Hi-C. For example, HiChIP achieves this by using chromatin immunoprecipitation to capture the genomic regions associated with a specific histone modification or protein of interest after Hi-C library preparation (Mumbach et al., 2017, 2016), whereas capture Hi-C uses oligonucleotide baits to capture genomic regions of interest (Dryden et al., 2014). Previous studies have used these techniques to identify causal genes at GWAS loci (Cairns et al., 2016; Dryden et al., 2014; Jäger et al., 2015; Martin et al., 2016, 2015; McGovern et al., 2016; Mumbach et al., 2017), uncovering many important mechanisms and pathways involved in various diseases. These studies have, however, mainly focused on cells derived from blood and immune cells.

Multiple publications have shown that chromatin interactions are cell-type specific and are altered during differentiation and stimulation (Burren et al., 2017; Dixon et al.,

¹Centre 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, Manchester, United Kingdom; ²Dermatology Centre, Salford Royal NHS Foundation Trust, NIHR Manchester Biomedical Research Centre, Manchester Academic Health Science Centre, Manchester, United Kingdom; ³Genes & Human Disease Research Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, USA; ⁴Division of Infection, Immunity and Respiratory Medicine, School of Biological Sciences, University of Manchester, Manchester, United Kingdom; ⁵Lydia Becker Institute of Immunology and Inflammation, The University of Manchester, Manchester, United Kingdom; ⁶NIHR Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, United Kingdom; and ⁷Division of Informatics, Imaging and Data Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, United Kingdom

⁸These authors contributed equally to this work

Correspondence: Chenfu Shi, 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, Room 1.023, AV Hill Building, Upper Brook Street, Manchester M13 9PL, United Kingdom. E-mail: chenfu.shi@postgrad.manchester.ac.uk

Abbreviations: AD, atopic dermatitis; eQTL, expression quantitative trait loci; kb, kilobase; KC, keratinocyte; Ps, psoriasis; PsA, psoriatic arthritis; TF, transcription factor

Received 2 June 2020; revised 12 January 2021; accepted 21 January 2021; accepted manuscript published online 17 February 2021; corrected proof published online 16 March 2021

¹Vösa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, et al. Unraveling the polygenic architecture of complex traits using blood eQTL meta-analysis. *bioRxiv* 2018. doi: <https://doi.org/10.1101/447367>.

2012; Hansen et al., 2018; Mumbach et al., 2017; Rao et al., 2014; Rubin et al., 2017; Schmitt et al., 2016; Siersbæk et al., 2017). Although for many autoimmune diseases associated risk variants primarily affect immune cells, other cell types might also be involved in the disease development (Farh et al., 2015; Mahajan et al., 2018; Mizoguchi et al., 2018). Autoimmune dermatological traits such as Ps, psoriatic arthritis (PsA), and systemic sclerosis are systemic conditions with heterogeneous effects in multiple cell types and are likely to involve a complex interplay between skin-resident cells and immune cells (Albanesi et al., 2018; McCoy et al., 2017).

Ps is recognized as an immune-mediated condition, and as such, the roles of immune cells in disease have received great attention. However, more recently, there has been a rejuvenated interest in the role of keratinocytes (KCs), which are the most predominant cells in the epidermis and are highly dysregulated in the disease in terms of proliferation and differentiation (Ni and Lai, 2020). KCs respond to T-cell signals by producing proinflammatory cytokines that further contribute to T-cell activation (Albanesi et al., 2018; Benhadou et al., 2018; Lorscheid et al., 2019). A prevalent T-cell-released signal in inflamed tissue is IFN- γ . In Ps, IFN- γ is released in the classical T helper type 1 pathway (Lowe et al., 2008) and accumulates in psoriatic lesions (Schlaak et al., 1994; Uyemura et al., 1993), promoting epidermal KC apoptosis (Hijnen et al., 2013).

Another recurrent feature in inflammatory skin diseases is the invasion of CD8+ T cells in the inflammation site (Antohe et al., 2019; Cai et al., 2012; Hennino et al., 2007; O'reilly et al., 2012). In Ps, these cells release IL-17, an important factor in disease pathology (Ortega et al., 2009).

In this study, we use HiChIP to map active chromatin interactions genome wide on spontaneously immortalized KCs (HaCaT), unstimulated or stimulated with IFN- γ , and a CD8+ T cell line derived from a cancerous skin plaque (MyLa). In contrast to capture Hi-C, this technique specifically enriches for active regions of the genome. Moreover, it provides interactions genome wide, allowing us to discover candidate genes for a larger set of disorders and to include more recently identified loci in our analysis. We use these new datasets to study disease-associated loci for Ps, PsA, AD, melanoma, and systemic sclerosis and to identify potentially associated genes. Our analysis pipeline is available at https://github.com/ChenfuShi/keratinocyte_gene_link and allows the possibility of using custom input variants (such as non-genome-wide significant loci or credible SNP sets).

We complement our dataset with matched RNA sequencing and Hi-C. 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) reprocessed with the latest computational tools to investigate the GWAS-associated loci, which might be mediated by these cell populations.

We show that the interacting genes enrich for pathways that are highly relevant to the underlying disease mechanisms. Our results can also inform the association of different genes to specific loci, and we provide examples of this for four distinct Ps loci. Therefore, our findings update our view

on the underlying disease mechanisms of these traits, facilitating future studies and drug discovery.

RESULTS

A compendium of activity and chromatin interactions in KCs

Chromatin interactions specific for active regulatory elements, such as enhancers and promoters, were identified using HiChIP targeting the active histone mark H3K27ac. Interactions and peaks are highly cell-type specific, as shown by clustering and principal component analysis (Figure 1a and b and Supplementary Figure S1a and b). We noted some batch effects between replicates that caused an imbalance in the number of detected HiChIP loops between replicates (HaCaT cells) and observed that more variance was explained by batch than by IFN- γ stimulation in HaCaT cells in the principal component analysis plots. Nevertheless, there is substantial overlap in the interaction called between replicates (Supplementary Figure S2 and Supplementary Table S1). As expected, H3K27ac peaks were strongly enriched in GWAS SNPs compared with the background in our data. Moreover, dermatological conditions such as the ones studied in this study possess a stronger enrichment in KCs than other diseases (Supplementary Figure S3).

We identified >50,000 significant interactions from each of our HiChIP datasets (summary statistics in Supplementary Tables S1 and S2). The median interaction distance was 250 kilobase (kb), which is consistent with the results derived from the public datasets (Figure 1c and Supplementary Figure S4a). Moreover, the vast majority (90%) of interactions connected the regions within topologically associating domains identified from our matched Hi-C datasets (Figure 1d). This is consistent with recent reports about topologically associating domains determining the scope of gene regulation (Rao et al., 2014; Rowley and Corces, 2018). We compared this new dataset with our previous study that used capture Hi-C in the same cell lines and targeted GWAS loci from a number of autoimmune disorders (Ray-Jones et al., 2020). We found significant overlap between capture Hi-C and HiChIP, although HiChIP identifies significantly more genes that are specifically active in these cell types (Figure 1c and d, Supplementary Figures S4b and S5a and b, and Supplementary Materials and Methods). Moreover, we recovered most of the interacting genes highlighted in the previous work, such as *IL23A*, *STAT3*, *B3GNT2*, *COMMD1*, *ERRF1*, and *SOX4* (Ray-Jones et al., 2020).

RNA sequencing for IFN- γ -stimulated HaCaT cells revealed 535 differentially expressed (447 upregulated and 88 downregulated) genes that were enriched for pathways related to IFN- γ stimulation (Supplementary Figure S5c). This stimulation has a significant impact on the number of interactions that originate from these genes: genes overexpressed on stimulation have 7.08 interactions per gene in stimulated cells compared with 4.48 in unstimulated cells (Supplementary Figure S5d) ($\times 1.58$ in these genes compared with $\times 1.18$ overall interactions). Although it seemed that these interactions linked the regions that have increased H3K27 acetylation (mean: $\times 1.3$, median: $\times 1.64$, in stimulated cells compared with that in unstimulated cells), the majority of these linked peaks already have significant levels

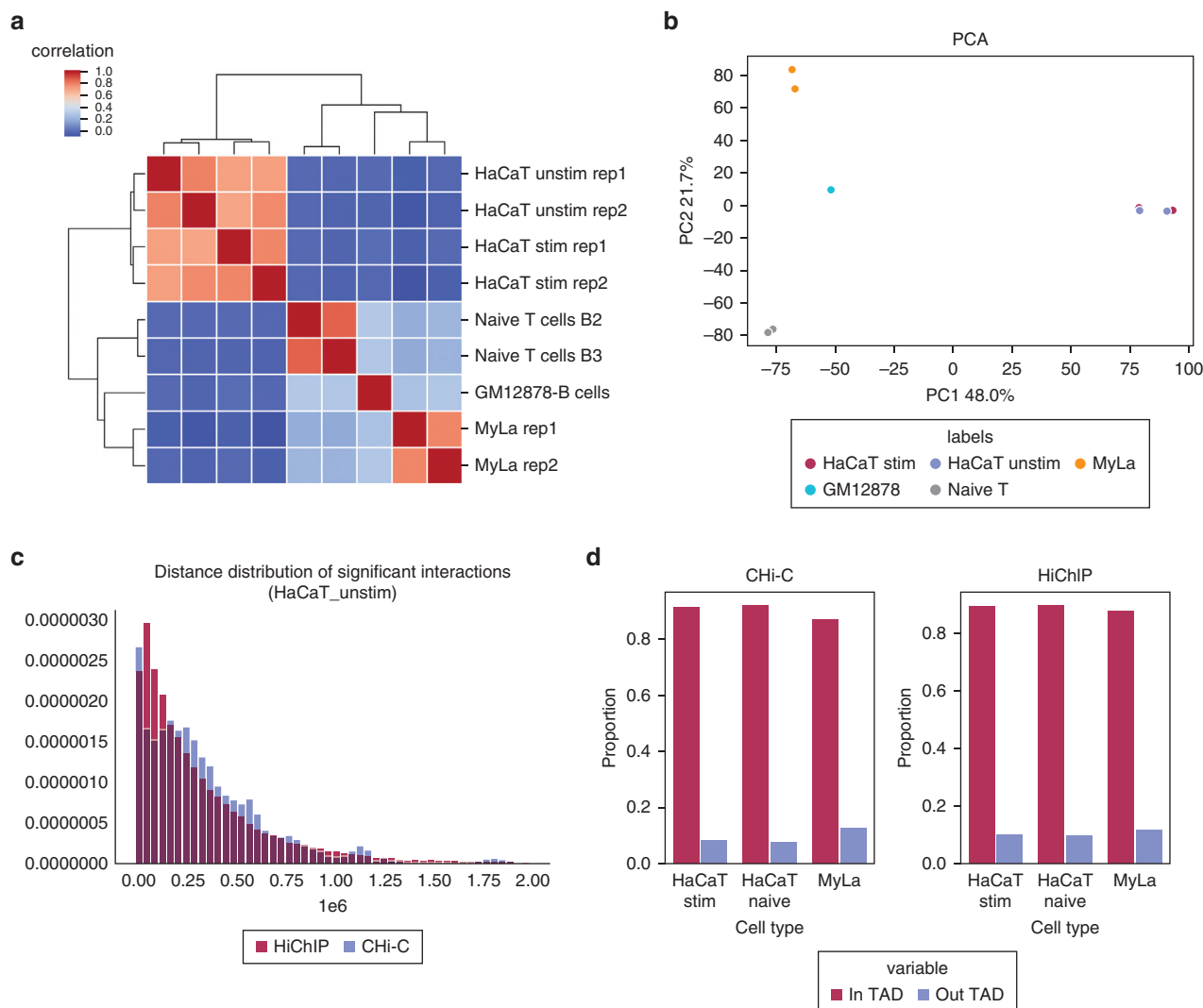


Figure 1. Properties of the chromatin looping datasets used in this study. (a) Hierarchical clustering of the loops for the individual HiChIP samples using correlation. (b) PCA of the loops for the individual HiChIP samples. (c) Comparison of the distance distribution of the significant fitchip (HiChIP) and CHiCAGO (CHI-C) interactions for naïve HaCaT HiChIP, a representative example. (d) The proportion of significant interactions that are within TADs and that cross TAD boundaries for region capture Hi-C data and HiChIP data. CHI-C, capture Hi-C; PC, principal component; PCA, principal component analysis; rep, replicate; stim, stimulated; TAD, topologically associating domain; unstim, unstimulated.

of H3K27 acetylation before stimulation (Supplementary Figure S5e).

Chromatin contacts identified by HiChIP pinpoint functionally relevant interactions

We first wanted to test whether HiChIP can successfully discover important features involved in gene regulation. To do this, we used the 535 differentially expressed genes that are activated on IFN- γ stimulation and assumed that these genes are regulated through enhancers binding transcription factors (TFs) related to IFN- γ response. As expected, H3K27ac peaks linked to these genes through HiChIP interactions in HaCaT were enriched for TF motifs that are known to be involved in the response to IFN- γ stimulation and/or viral infection (Ramezani et al., 2018; Schroder et al., 2004) and contained similar motifs to those identified in peaks that had different H3K27ac signal in the stimulated condition (Supplementary Figure S5f). H3K27ac peaks linked through HiChIP interactions to upregulated

genes specifically in stimulated cells enriched significantly for *IRF2*, *ISRE*, and *IRF1* motifs compared with the H3K27ac peaks linked to these genes in both conditions (Supplementary Table S3).

Next, we wanted to assess how HiChIP datasets can recapitulate results from large expression quantitative trait loci (eQTL) studies in the discovery of genes dysregulated by GWAS variants. Using Ps loci, we tested our results from naïve T cells and GM12878 cells with the largest blood eQTL database available, eQTLgen (Võsa et al., 2018¹), as ground truth. Our HiChIP analysis showed a 52% recall rate and 31% specificity compared with that of the genes identified from eQTLs. Using genes that were linked in both HiChIP replicates, we obtained a 46% recall rate and 40% specificity. For comparison, Javierre et al. (2016) reported a recall rate of 25.7% with lead eQTLs.

We then compared our HaCaT KC datasets with the Genotype-Tissue Expression dataset from sun-exposed skin

Table 1. Number of Genes Identified Associated with Each Disease Studied

Disease	Number of Genes Identified
PsA	84
Ps	399
Eczema	154
Melanoma	165
Scleroderma	182

Abbreviations: Ps, psoriasis; PsA, psoriasis arthritis.

(GTEx Consortium, 2013), resulting in a recall rate of 56% and a specificity of 14.7% (38% and 20%, respectively, with genes reproduced from both replicates). In this study, the Genotype-Tissue Expression dataset contains relatively few eQTLs compared with the eQTLgen database owing to the limited sample size, reducing specificity in our analysis. By comparison, our previous capture Hi-C dataset generated a 21% recall rate and 11% specificity. Interestingly, 38 of 52 eQTLs linked to Ps loci identified in the skin in Genotype-Tissue Expression are recapitulated in the eQTLgen dataset.

These eQTL datasets were generated from slightly different cell populations compared with those analyzed in this study. We have attempted using eQTL datasets generated from pure cell populations (for CD4+ and CD8+ T cells) such as the Kasela et al. (2017) dataset and the Database of Immune Cell Expression eQTL dataset (Schmiedel et al., 2018), but these studies produced very few hits within the studied loci owing to their limited sample size. Regardless, we obtained similar recall rates as reported earlier (22–66%).

Using chromatin conformation and functional genomics to dissect disease-associated loci

We next sought to use our dataset to identify the genes whose transcription start sites either overlap or are linked by chromatin interaction to GWAS SNPs associated with PsA, Ps, AD, melanoma, and systemic sclerosis. We identified between 84 and 399 potentially linked genes for each disease studied in our combined datasets (Table 1), with a large fraction of these being recoverable using both replicates for each cell type (Supplementary Figure S6). 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, whereas genes linked to Ps and PsA were linked to cytokine signaling and IL-23 (Figure 2), which is a pathway targeted by multiple novel treatments (Sakkas et al., 2019).

Although some genes were common in all the cell lines (e.g., 12% in Ps), most genes were specific to one or a few cell types (Supplementary Figure S7). Moreover, we found that most loci also implicate multiple genes, with an average of 7.1 genes per loci in Ps. Interestingly, the number of genes implicated by our method correlated significantly with the number of genes that were linked by eQTLs for the same loci ($P = 4.61e-13$) (Supplementary Figure S8). This could be partly caused by the size of the linkage disequilibrium block of the loci, in which the lead SNP implicates a set of co-inherited SNPs that can have effects on the expression of many genes.

Using these data, we can begin to provide some functional insight into the mechanisms mediating disease susceptibility. We found that for a significant number of loci, the closest protein-coding genes were not implicated, despite it being possible for linkage disequilibrium blocks to have multiple closest genes. For example, of the 44 Ps loci tested with a linkage disequilibrium block smaller than 100 kb, we linked all the closest genes for only 25 loci. For 14 loci, we did not find any evidence of the closest genes being regulated by the GWAS SNPs in any of the cell lines tested, and for 5 loci, we linked some but not all of the closest genes. Similar numbers can be seen in all the traits studied (Supplementary Table S4). We also identified some loci that do not have any protein-coding genes within 50 kb. We identified 12 of these loci across the diseases studied and provided putative linked genes for all of them in Supplementary File S1.

Looking at a locus, particularly the *ETS1* locus, Ps and AD have distinct associations. However, whereas the Ps locus directly overlaps the promoter of *ETS1*, the association with AD is located 130 kb downstream of the gene. We observed significant chromatin interactions between the AD locus and the *ETS1* promoter in naïve T, MyLa, and GM12878 cells (Supplementary Figures S9 and S10). Our data suggest a putative mechanism in which the distinct disease associations at this locus are mediated by a single gene. We follow up with the description of other examples in Ps in the next section.

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

The results from our analysis can be applied to augment the existing understanding of disease-associated loci and to link new genes or change the ones currently linked. In this study, we focus on four Ps-associated loci that show cell-type-specific and, to our knowledge, previously unreported interactions.

The Ps locus indexed by SNP rs73178598 is located in an intergenic region overlapping an antisense RNA, *SATB1-AS1*. We found a 240 kb T-cell-specific interaction present in naïve T and MyLa cells linking this locus with the promoter of *SATB1* in all replicates (Figure 3). Interestingly, for this region, there is an interaction that is present in naïve T cells and MyLa but not in KCs or GM12878 with differences visible in Hi-C maps as well (Supplementary Figure S11). Silencing of *SATB1* has been shown to have a similar effect to IFN- γ stimulation on major histocompatibility complex chromatin organization (Kumar et al., 2007) and is an important regulator of regulatory T cells and autoimmunity (Beyer et al., 2011).

The Ps loci indexed by SNP rs9504361 is intronic to *EXOC2* and is typically associated with this gene. However, H3K27ac occupancy at the locus is specific to MyLa cells in our analysis. We detected long-range interactions between this SNP and the promoters of *IRF4* (all replicates) and *DUSP22* (one replicate only), specifically in this cell type (Supplementary Figures S12 and S13). *IRF4* is a lymphocyte-specific TF central to the activation of innate and adaptive immune systems (Huber and Lohoff, 2014), and it was found to be overexpressed in psoriatic skin lesions (Ni et al., 2012). *DUSP22* is a phosphatase that might be involved in the c-Jun N-terminal kinase signaling pathway and has been shown to

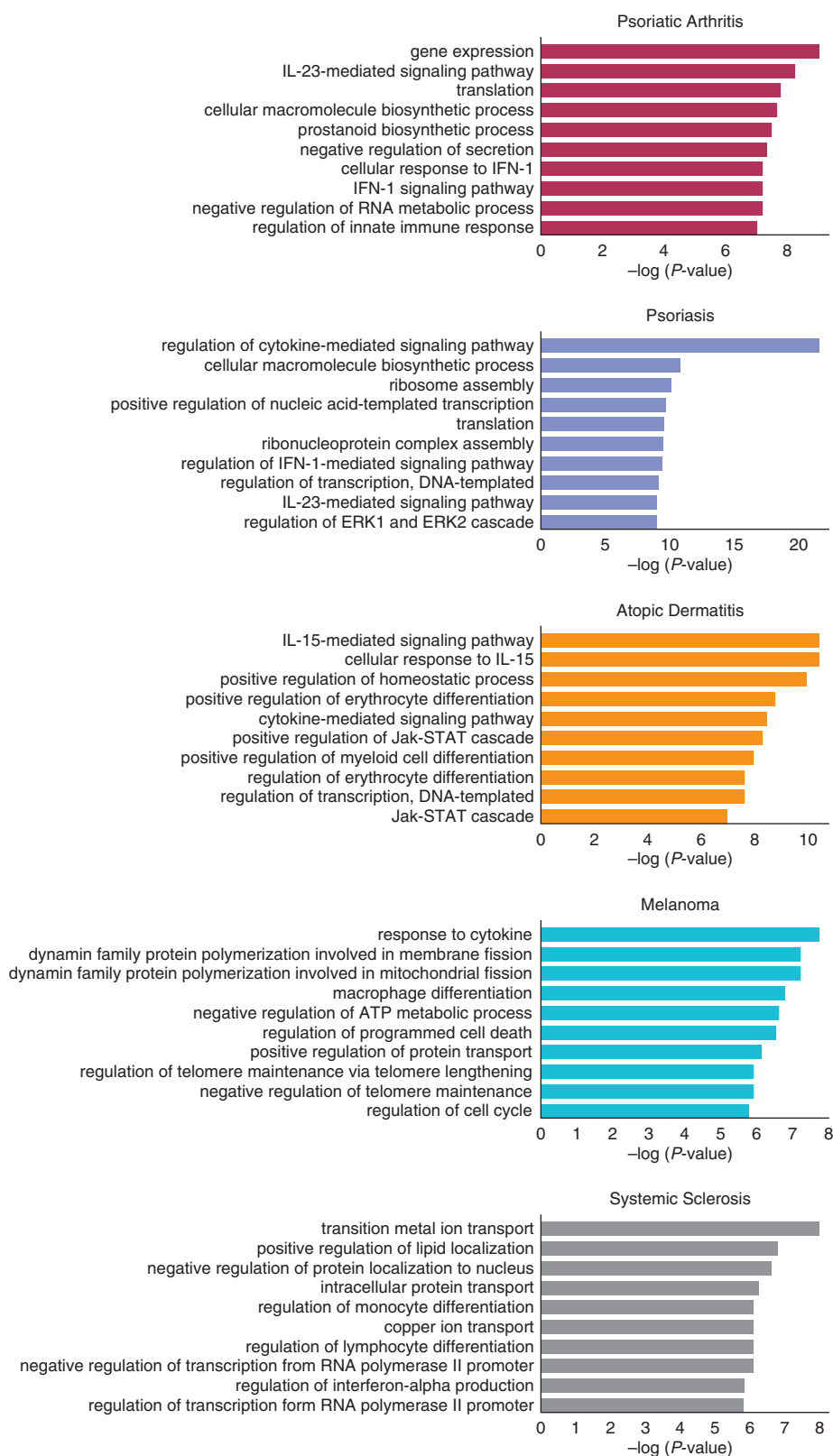


Figure 2. The top pathways enriched by the genes linked to disease-associated loci from all cell types.

Pathways were identified by EnrichR using genes that were linked in both technical replicates for each condition and/or cell type. ATP, adenosine triphosphate; ERK, extracellular signal-regulated kinase; STAT, signal transducer and activator of transcription.

be associated with lymphomas (Paydas et al., 2019; Zeke et al., 2016).

Another example of a gene target, to our knowledge previously unreported, can be found at the 1p36 locus (rs10794648), which to date has been associated with

IFNLR1 (also known as *IL28-RA*) because it is the closest gene to the associated SNPs (Genetic Analysis of Psoriasis Consortium & the Wellcome Trust Case Control Consortium 2, et al., 2010; Stuart et al., 2015). *IFNLR1* encodes part of a receptor for IFN- γ that is presented in the epidermis and is

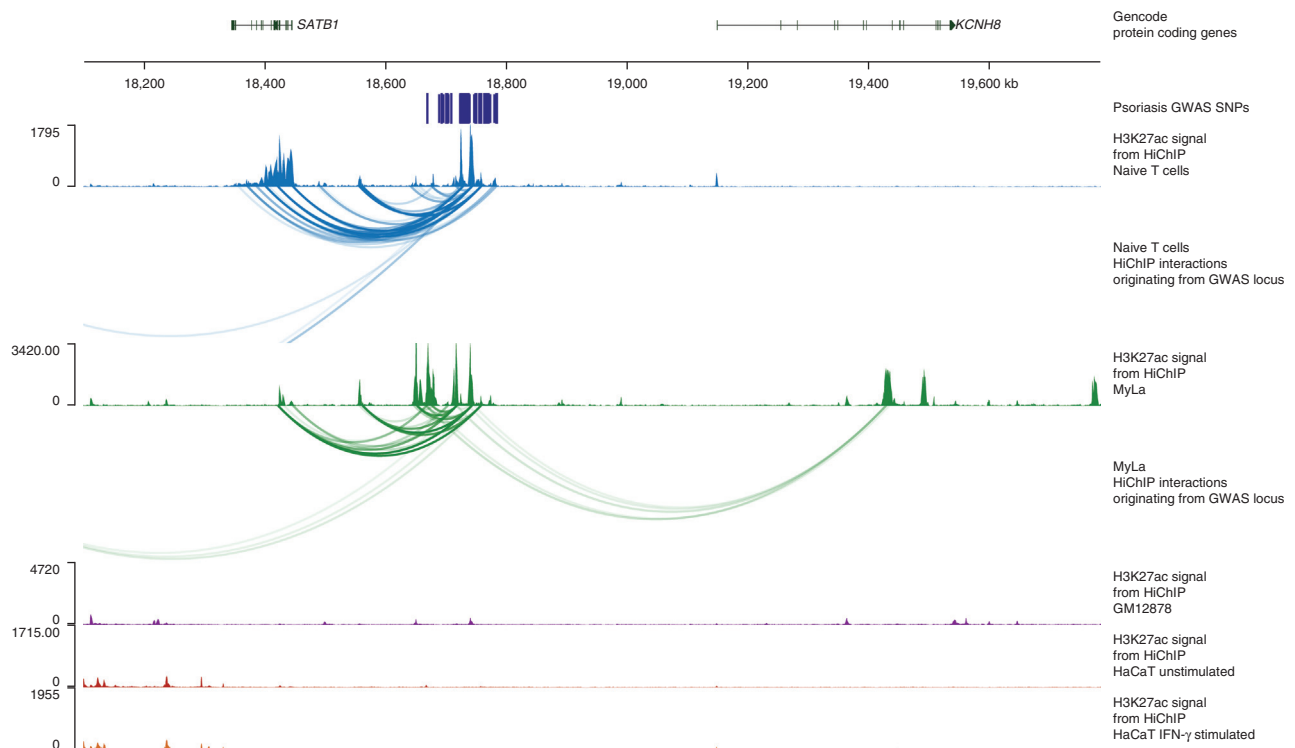


Figure 3. HiChIP interactions from the *SATB1/KCNH8* locus link *SATB1* as the target gene in this locus. The tracks (in order) include the following: gene code genes, SNPs associated with psoriasis ($r^2 > 0.8$), H3K27ac signal in naive T cells, significant long-range interactions originating from the psoriasis-associated locus in naive T cells, H3K27ac signal in MyLa cells, significant long-range interactions originating from the psoriasis associated locus in MyLa cells, H3K27ac signal in GM12878 cells, H3K27ac signal in IFN- γ -stimulated HaCaT cells, and H3K27ac signal in naive T cells. kb, kilobase.

thought to promote an antiviral response in Ps (Lazear et al., 2015). However, our HiChIP data showed long-range interactions between the Ps SNPs at 1p36 and the distal gene *GRHL3* primarily in HaCaT KCs (all replicates in unstimulated cells and combined only in IFN- γ stimulated cells) (Figure 4a and Supplementary Figures S14 and S15). We further explored the chromatin profile in this region using available chromatin immunoprecipitation sequencing data for enhancer-associated marks (H3K27ac and H3K4me1) and a promoter-associated mark (H3K4me3) in progenitor, differentiating and migrating primary KCs (Klein et al., 2017). We found that *GRHL3* overlaps H3K27ac and H3K4me3 marks in all conditions, indicating that the promoter is active through differentiation (Supplementary Figure S16a). Moreover, we found that the Ps-associated SNPs overlap H3K4me1 in migrating and H3K27ac in migrating and progenitor cells but not H3K4me3 in any condition (Supplementary Figure S16b). These findings suggest that the SNPs overlap an enhancer that is present in migrating and progenitor cells but are active in progenitor cells only. *GRHL3* encodes a TF that stimulates migration. It is upregulated in psoriatic lesions and is required for repair of the epidermal skin barrier after immune-mediated injury (Gordon et al., 2014). Improper regulation of *GRHL3* in progenitor cells could adversely affect the migration of KCs, which is known to be highly accelerated in Ps, leading to the formation of the classic skin plaques. Targets of the *GRHL3* TF include further GWAS-implicated genes such as *IVL* (involved in KC differentiation) (Watt, 1983) and *KLF4* (TF involved in KC differentiation and skin barrier formation). *KLF4* was recently implicated as the likely

functional target gene in the 9q31 locus in our previous study (Ray-Jones et al., 2020).

Finally, the locus indexed by the Ps SNP rs73183592, to our knowledge, previously undescribed mechanism was linked through a long-range interaction spanning about 500 kb to *FOXO1* (Figure 4b and Supplementary Figure S17), a gene with important functions in regulatory T cells. Interestingly, this interaction was identified primarily in KCs in our analysis (all replicates in unstimulated and one replicate only in IFN- γ stimulated) but was only weakly supported in T cells because this enhancer seems to be specific to KCs. Dysregulation of this pathway was also found to be important in the development of Ps (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 (Burren et al., 2017; Dixon et al., 2012; Hansen et al., 2018; Mumbach et al., 2017; Rao et al., 2014; Rubin et al., 2017; Schmitt et al., 2016; Siersbæk et al., 2017).

In this study, we used H3K27ac HiChIP, a modern technique that allows combined analysis of both chromatin conformation and chromatin activity, to create a global study of promoter–enhancer interactions in KC and CD8+ T-cell lines. Equivalent data have so far only been generated in

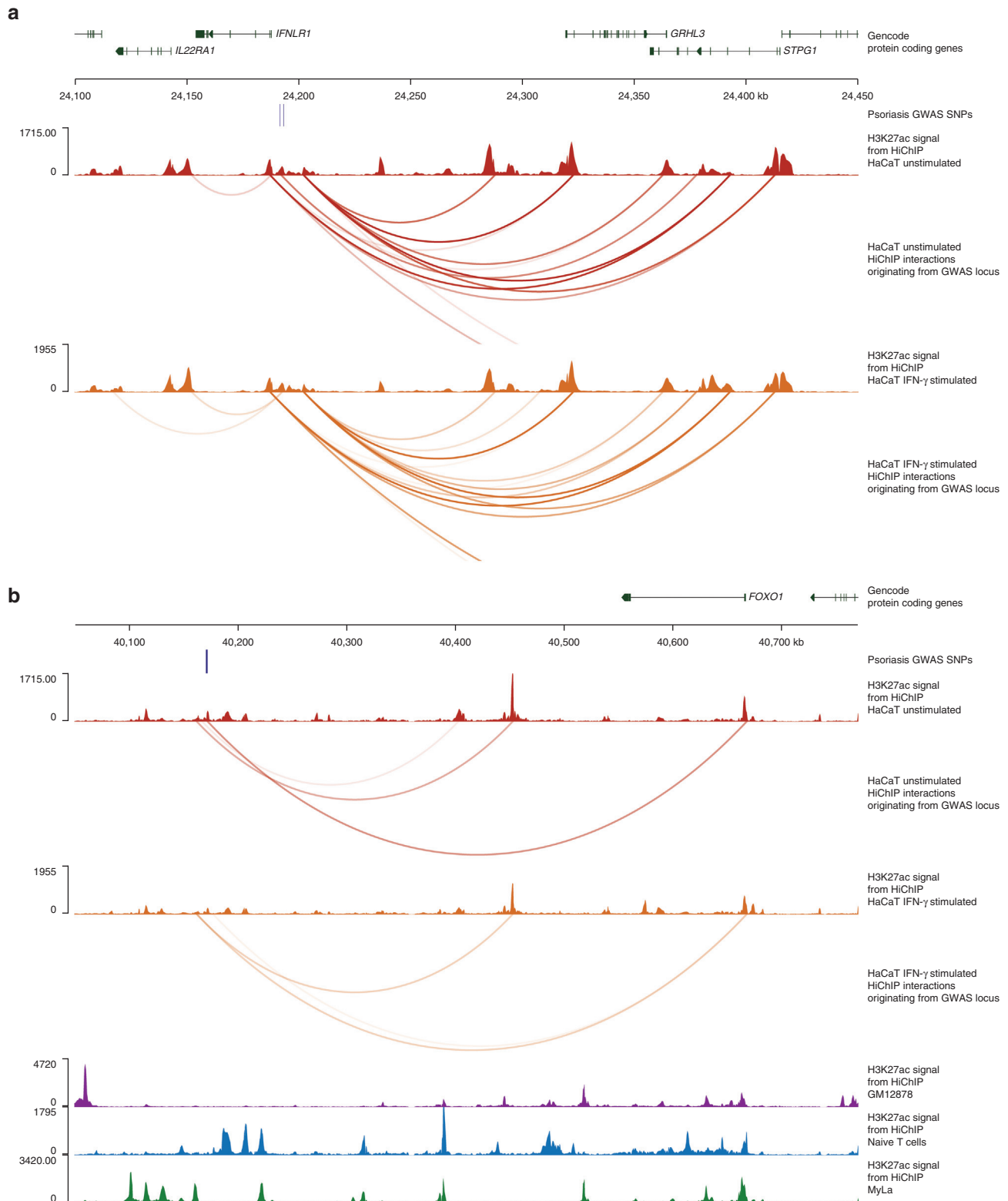


Figure 4. HiChIP interactions link *GRHL3* and *FOXO1* as candidate genes in psoriasis. (a) HiChIP interactions from the *IFNL1/GRHL3* locus link *GRHL3* as a candidate gene. The tracks (in order) include the following: gencode genes, SNPs associated with psoriasis ($r^2 > 0.8$), H3K27ac signal in unstimulated HaCaT cells, significant long-range interactions originating from the psoriasis-associated locus in unstimulated HaCaT cells, H3K27ac signal in IFN- γ -stimulated HaCaT cells, and significant long-range interactions originating from the psoriasis-associated locus in IFN- γ -stimulated HaCaT cells. (b) HiChIP interactions from the *FOXO1* locus link *FOXO1* as a candidate gene. The tracks (in order) include the following: gencode genes, SNPs associated with psoriasis ($r^2 > 0.8$), H3K27ac signal in unstimulated HaCaT cells, significant long-range interactions originating from the psoriasis-associated locus in unstimulated HaCaT cells, H3K27ac signal in IFN- γ -stimulated HaCaT cells, significant long-range interactions originating from the psoriasis-associated locus in IFN- γ -stimulated HaCaT cells, H3K27ac signal in GM12878 cells, H3K27ac signal in naive T cells, and H3K27ac signal in MyLa cells. kb, kilobase.

immune cells with few examples in other cell populations, such as HiChIP in endometrial cancer cells (O'Mara et al., 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).

After assessing the effectiveness of these techniques in linking functional elements with genes in a cell-type-specific manner, we show their possible use in studying disease-associated loci. We explore disease-associated SNPs for PsA, Ps, AD, melanoma, and systemic sclerosis and identify all the genes that are linked by chromatin interactions to these variants. We show that these genes are enriched for disease-relevant pathways and provide tables and figures for all the loci (Supplementary File S1 and GitHub repository). We show how these data, using four distinct Ps-associated loci as examples, allow us to identify previously unreported mechanisms and provide functional insight into the diseases studied.

The HiChIP gene targets 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). As a proof of concept, we tested the genes identified in this work by querying the DrugBank database (Wishart et al., 2008) in search of drugs that are currently used and the ones that could be repurposed. Across the diseases studied, we identified 127 genes that are targeted by approved drugs in some diseases, corresponding to 231 drugs that could be potentially repurposed (Supplementary File S2).

A limitation of this study is that it is based on immortalized cell lines. In the future, these studies should include primary tissues and compare healthy with diseased samples. This has so far been done in a limited way, but early results are showing a significant impact of disease state and genetics on chromatin conformation (Delaneau et al., 2019; Fasolino et al., 2020; Gorkin et al., 2019; Kloetgen et al., 2020). Moreover, although the techniques used in this study are an important strategy to identify potential candidate genes, they require further individual experimental validation (such as genetic perturbation techniques) before they can be used for drug targeting.

In summary, through our analysis, we present a list of potential target genes and pathways for mediating disease risk for five complex diseases, and by documenting individual loci, we highlight some mechanisms by which this risk is mediated. These genes and mechanisms represent a useful resource for further research aimed at characterizing how genetic variation has an impact on disease susceptibility for these and other complex diseases.

MATERIALS AND METHODS

For detailed methods, see the [Supplementary Materials and Methods](#).

Cell culture

HaCaT KC cells (T0020001, Addexbio Technologies, San Diego, CA) were cultured in high-glucose DMEM supplemented with 10% fetal bovine serum and penicillin-streptomycin (final concentration: 100 U penicillin and 0.1 mg/ml streptomycin). For HaCaT stimulation experiments, the media were supplemented with 100 ng/ml

recombinant human IFN- γ (285-IF-100; R&D Systems, Minneapolis, MN), and cells were incubated for 8 hours before harvest.

My-La CD8+ cells (95051033, Sigma-Aldrich, St. Louis, MO) were cultured in RPMI 1640 medium supplemented with 10% AB human serum (Sigma Aldrich), 100 U/ml recombinant human IL-2 (Sigma-Aldrich), and penicillin-streptomycin.

HiChIP experiments

HiChIP libraries were generated according to the Chang Lab protocol (Mumbach et al., 2016). Briefly, 10 million cross-linked cells were lysed, the chromatin was digested with MboI, biotinylated, and ligated. After random shearing, chromatin immunoprecipitation was performed against H3K27ac (ab4729, Abcam, Cambridge, United Kingdom). Decrosslinking and biotin pull-down preceded tagmentation with Tn5 transposase (Illumina, San Diego, CA) at 55 °C for 10 minutes. For more details, see the [Supplementary Materials and Methods](#).

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 a transcription start site was located within 5 kb of a loop. We also associated the transcripts for which the transcription start site was within 1 kb of an SNP overlapping an H3K27ac peak as identified from HiChIP data. All transcripts were then grouped by gene, and the genes were filtered by an expression level of at least 1 transcript per million in the corresponding cell line.

Data availability statement

The sequence datasets generated and analyzed during this study are available in the Gene Expression Omnibus repository under accession numbers GSE137906 (for capture Hi-C only) and GSE151193 (for HiChIP, RNA sequencing, and Hi-C with updated processed files).

The code required to reproduce the analysis in this study is available on GitHub. Figures with Hi-C and HiChIP loops for every locus for all the diseases studied in this research are also available in the same repository (https://github.com/ChenfuShi/keratinocyte_gene_link).

ORCIDiDs

Chenfu Shi: <http://orcid.org/0000-0002-3083-8416>
 Helen Ray-Jones: <http://orcid.org/0000-0002-8884-6865>
 James Ding: <http://orcid.org/0000-0001-7273-9646>
 Kate Duffus: <http://orcid.org/0000-0002-8925-1863>
 Yao Fu: <http://orcid.org/0000-0003-2348-4872>
 Vasanthi Priyadarshini Gaddi: <http://orcid.org/0000-0001-8642-6907>
 Oliver Gough: <http://orcid.org/0000-0002-4737-1023>
 Jenny Hankinson: <http://orcid.org/0000-0002-9021-4031>
 Paul Martin: <http://orcid.org/0000-0002-1016-6851>
 Amanda McGovern: <http://orcid.org/0000-0001-7727-3283>
 Annie Yarwood: <http://orcid.org/0000-0001-5026-9815>
 Patrick Gaffney: <http://orcid.org/0000-0002-1580-869X>
 Steve Eyre: <http://orcid.org/0000-0002-1251-6974>
 Magnus Rattray: <http://orcid.org/0000-0001-8196-5565>
 Richard B Warren: <http://orcid.org/0000-0002-2918-6481>
 Gisela Orozco: <http://orcid.org/0000-0002-3479-0448>

CONFLICT OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGMENTS

The authors would like to acknowledge the assistance given by Information Technology Services and the use of the Computational Shared Facility at The University of Manchester, United Kingdom. This work was funded by the

Wellcome Trust (award references 207491/Z/17/Z and 215207/Z/19/Z), the Versus Arthritis (award reference 21754), the NIHR Manchester Biomedical Research Centre, and the Medical Research Council (award reference MR/N00017X/1).

AUTHOR CONTRIBUTIONS

Conceptualization: SE, RBW, GO, MR, CS, HRJ; Data curation: CS, HRJ, PM, VPG; Formal Analysis: CS; Funding Acquisition: GO, MR, RBW, SE; Investigation: CS, HRJ, JH; Methodology: OG, JH, PM, AMG, AY, PG, YF, KD, JD; Project Administration: RBW, SE, GO; Resources: VPG, PM, PG, YF, GO, SE, MR, RBW; Software: CS; Supervision: GO, MR, SE, RBW; Writing – Original Draft Preparation: CS, HRJ, JD; Writing – Review and Editing: CS, HRJ, JD, KD, YF, VPG, OG, JH, PM, AM, AY, PG, SE, MR, RBW, GO

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at <https://doi.org/10.1016/j.jid.2021.01.015>.

REFERENCES

- Albanesi C, Madonna S, Gisondi P, Girolomoni G. The interplay between keratinocytes and immune cells in the pathogenesis of psoriasis. *Front Immunol* 2018;9:1549.
- Antohe M, Nedelcu RI, Nichita L, Popp CG, Cioplea M, Brinzea A, et al. Tumor infiltrating lymphocytes: the regulator of melanoma evolution. *Oncol Lett* 2019;17:4155–61.
- Benhadou F, Mintoff D, Schnebert B, Thio HB. Psoriasis and microbiota: a systematic review. *Diseases* 2018;6:47.
- Beyer M, Thabet Y, Müller RU, Sadlon T, Classen S, Lahl K, 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.
- Burren OS, Rubio García A, Javierre BM, Rainbow DB, Cairns J, Cooper NJ, 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, Várnai C, Dimond A, Plagnol V, et al. CHiCAGO: robust detection of DNA looping interactions in capture Hi-C data. *Genome Biol* 2016;17:127.
- Choy MK, Javierre BM, Williams SG, Baross SL, Liu Y, Wingett SW, et al. Promoter interactome of human embryonic stem cell-derived cardiomyocytes connects GWAS regions to cardiac gene networks [published correction appears in *Nat Commun* 2018;12:9:4792]. *Nat Commun* 2018;9:2526.
- Delaneau O, Zazhytska M, Borel C, Giannuzzi G, Rey G, Howald C, et al. Chromatin three-dimensional interactions mediate genetic effects on gene expression. *Science* 2019;364:eaat8266.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 2012;485:376–80.
- Dryden NH, Broome LR, Dudbridge F, Johnson N, Orr N, Schoenfelder S, et al. Unbiased analysis of potential targets of breast cancer susceptibility loci by capture Hi-C. *Genome Res* 2014;24:1854–68.
- Ernst J, Kheradpour P, Mikkelsen TS, Shores N, Ward LD, Epstein CB, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 2011;473:43–9.
- Fang H, , ULTRA-DD Consortium, De Wolf H, Knezevic B, Burnham KL, Osgood J, et al. A genetics-led approach defines the drug target landscape of 30 immune-related traits. *Nat Genet* 2019;51:1082–91.
- Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* 2015;518:337–43.
- Fasolino M, Goldman N, Wang W, Cattau B, Zhou Y, Petrovic J, et al. Genetic variation in Type 1 diabetes reconfigures the 3D chromatin organization of T cells and alters gene expression. *Immunity* 2020;52:257–74.e11.
- Genetic Analysis of Psoriasis Consortium & the Wellcome Trust Case Control Consortium 2, Strange A, Capon F, Spencer CC, Knight J, Weale ME, et al. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nat Genet* 2010;42:985–90.
- Gordon WM, Zeller MD, Klein RH, Swindell WR, Ho H, Espetia F, et al. A GRHL3-regulated repair pathway suppresses immune-mediated epidermal hyperplasia. *J Clin Invest* 2014;124:5205–18.
- Gorkin DU, Qiu Y, Hu M, Fletez-Brant K, Liu T, Schmitt AD, et al. Common DNA sequence variation influences 3-dimensional conformation of the human genome. *Genome Biol* 2019;20:255.
- GTEX Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45:580–5.
- GTEx Consortium, Laboratory. Data Analysis & Coordinating Center (LDACC)—Analysis Working Group, Statistical Methods Groups—Analysis Working Group, Enhancing GTEx (eGTEx) Groups, NIH Common Fund, NIH/NCI, et al. Genetic effects on gene expression across human tissues [published correction appears in *Nature* 2018;553:530]. *Nature* 2017;550:204–13.
- Hansen AS, Cattoglio C, Darzacq X, Tjian R. Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus* 2018;9:20–32.
- Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for the ENCODE project. *Genome Res* 2012;22:1760–74.
- Hennino A, Vocanson M, Toussaint Y, Rodet K, Benetière J, Schmitt AM, et al. Skin-infiltrating CD8 + T cells initiate atopic dermatitis lesions. *J Immunol* 2007;178:5571–7.
- Hijnen D, Knol EF, Gent YY, Giovannone B, Beijin SJ, Kupper TS, 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, Miglioni G, Henrion M, Kandaswamy R, Speedy HE, Heindl A, 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, Kreuzhuber R, Hill SM, Sewitz S, et al. Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. *Cell* 2016;167:1369–84.e19.
- Kasela S, Kisand K, Tserel L, Kaleviste E, Remm A, Fischer K, et al. Pathogenic implications for autoimmune mechanisms derived by comparative eQTL analysis of CD4+ versus CD8+ T cells. *PLoS Genet* 2017;13:e1006643.
- Klein RH, Lin Z, Hopkin AS, Gordon W, Tsoi LC, Liang Y, et al. GRHL3 binding and enhancers rearrange as epidermal keratinocytes transition between functional states. *PLoS Genet* 2017;13:e1006745.
- Kloetgen A, Thandapani P, Ntziachristos P, Ghebrechristos Y, Nomikou S, Lazaris C, et al. Three-dimensional chromatin landscapes in T cell acute lymphoblastic leukemia. *Nat Genet* 2020;52:388–400.
- Kumar PP, Bischof O, Purbey PK, Notani D, Urlaub H, Dejean A, 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.
- Lazear HM, Nice TJ, Diamond MS. Interferon- λ : immune functions at barrier surfaces and beyond. *Immunity* 2015;43:15–28.
- Li B, Lei J, Yang L, Gao C, Dang E, Cao T, 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.
- Lorscheid S, Müller A, Löffler J, Resch C, Bucher P, Kurschus FC, et al. Keratinocyte-derived I κ B ζ drives psoriasis and associated systemic inflammation. *JCI Insight* 2019;4:e130835.
- Lowes MA, Kikuchi T, Fuentes-Duculan J, Cardinale I, Zaba LC, Haider AS, 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, Robertson NR, Torres JM, Rayner NW, 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 P, Ding J, Duffus K, Gaddi VP, McGovern A, Ray-Jones H, 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, Schoenfelder S, Duffus K, Yarwood A, 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, Duffus K, Yarwood A, Schoenfelder S, et al. Capture Hi-C reveals novel candidate genes and complex long-range interactions with related autoimmune risk loci. *Nat Commun* 2015;6:10069.
- Mccooy SS, Reed TJ, Berthier CC, Tsou PS, Liu J, Gudjonsson JE, et al. Scleroderma keratinocytes promote fibroblast activation independent of transforming growth factor beta. *Rheumatology (Oxford)* 2017;56:1970–81.
- McGovern A, Schoenfelder S, Martin P, Massey J, Duffus K, Plant D, et al. Capture Hi-C identifies a novel causal gene, IL20RA, in the pan-autoimmune genetic susceptibility region 6q23. *Genome Biol* 2016;17:212.
- Miguel-Escalada I, Bonàs-Guarch S, Cebola I, Ponsa-Cobas J, Mendieta-Esteban J, Atla G, 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, Marshall JL, Rao DA, Chang SK, et al. Functionally distinct disease-associated fibroblast subsets in rheumatoid arthritis. *Nat Commun* 2018;9:789.
- Mumbach MR, Rubin AJ, Flynn RA, Dai C, Khavari PA, Greenleaf WJ, et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods* 2016;13:919–22.
- Mumbach MR, Satpathy AT, Boyle EA, Dai C, Gowen BC, Cho SW, 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, Li W, Chen S, Li J. 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.
- Ni X, Lai Y. Keratinocyte: a trigger or an executor of psoriasis? *J Leukoc Biol* 2020;108:485–91.
- Nica AC, Dermizakis 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, Endometrial Cancer Association Consortium. 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, Hügler T, Van Laar JM. T cells in systemic sclerosis: a reappraisal. *Rheumatology (Oxford)* 2012;51:1540–9.
- Ortega C, Fernández-AS, Carrillo JM, Romero P, Molina JJ, Moreno JC, et al. IL-17-producing CD8 + T lymphocytes from psoriasis skin plaques are cytotoxic effector cells that secrete Th17-related cytokines. *J Leukoc Biol* 2009;86:435–43.
- Paternoster L, Standl M, Waage J, Baurecht H, Hotze M, Strachan DP, 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.
- Paydas S, Bagir EK, Ergin M, Seydaoglu G, Boz A. The role of DUSP22 (dual specificity phosphatase 22) gene expression in the prognosis of low grade lymphomas. *Hematol Oncol* 2019;37:340–1.
- Ramezani A, Nahad MP, Faghihloo E. The role of Nrf2 transcription factor in viral infection. *J Cell Biochem* 2018;119:6366–82.
- Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, 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, Martin P, Shi C, Hankinson J, et al. Mapping DNA interaction landscapes in psoriasis susceptibility loci highlights KLF4 as a target gene in 9q31. *BMC Biol* 2020;18:47.
- Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. *Nature* 2015;518:317–30.
- Rowley MJ, Corces VG. Organizational principles of 3D genome architecture. *Nat Rev Genet* 2018;19:789–800.
- Rubin AJ, Barajas BC, Furlan-Magaril M, Lopez-Pajares V, Mumbach MR, Howard I, 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, Hermann E, Girndt M, Gallati H, et al. T cells involved in psoriasis vulgaris belong to the Th1 subset. *J Invest Dermatol* 1994;102:145–9.
- Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, et al. Impact of genetic polymorphisms on human immune cell gene expression. *Cell* 2018;175:1701–15.e16.
- Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan CL, et al. A compendium of chromatin contact maps reveals spatially active regions in the human genome. *Cell Rep* 2016;17:2042–59.
- Schroder K, Hertzog PJ, Ravasi T, Hume DA. Interferon-gamma: an overview of signals, mechanisms and functions. *J Leukoc Biol* 2004;75:163–89.
- Siersbæk R, Madsen JGS, Javierre BM, Nielsen R, Bagge EK, Cairns J, et al. Dynamic rewiring of promoter-anchored chromatin loops during adipocyte differentiation. *Mol Cell* 2017;66:420–35.e5.
- Song M, Yang X, Ren X, Maliskova L, Li B, Jones IR, et al. Mapping cis-regulatory chromatin contacts in neural cells links neuropsychiatric disorder risk variants to target genes. *Nat Genet* 2019;51:1252–62.
- Stuart PE, Nair RP, Tsoi LC, Tejasvi T, Das S, Kang HM, 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, Gudjonsson JE, Das S, Zawistowski M, 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, Modlin RL, Nickoloff BJ. 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.
- Watt FM. Involucrin and other markers of keratinocyte terminal differentiation. *J Invest Dermatol* 1983;81(Suppl). 100s–3s.
- Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, Tzur D, et al. DrugBank: a KnowledgeBase for drugs, drug actions and drug targets. *Nucleic Acids Res* 2008;36:D901–6.
- Zeke A, Misheva M, Reményi A, Bogoyevitch MA. 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.



This work is licensed under a Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>