Genetics of Scarring

Ung, Chuin Ying

Awarding institution:
King's College London

Genetics of Scarring

Chuin Ying Ung

Genetics & Molecular Medicine Research Division

King's College London

A thesis submitted for the degree of

Doctor of Philosophy

September 2022

For Alasdair Warwick

Acknowledgements

Funding: I received funding from the Medical Research Council (MR/S000348/1).

Contribution by other researchers: The keloid/excessive scarring pedigree was identified by Professor Eli Sprecher (Tel-Aviv Sourasky Medical Centre) in collaboration with Professor John McGrath (St John's Institute of Dermatology, King's College London). Pedigree samples were submitted for whole exome sequencing and validated with Sanger sequencing by Dr Alexandros Onoufriadis and William Scott (St John's Institute of Dermatology, King's College London), with whom I worked closely to analyse the whole exome sequencing data.

Control primary dermal fibroblasts and skin samples were provided by Dr Tanya Shaw (Centre for Inflammation Biology and Cancer Immunology, King's College London; ethics reference 14/NS/1073). Primary dermal fibroblasts carrying the homozygous *C1QTNF12* c.121dupC variant were gifted by Professor Eli Sprecher and Dr Ofer Sarig (ethics reference 0536-08-TLV, Tel-Aviv Sourasky Medical Center). Adipogenic differentiation of dermal fibroblasts were performed by Dr Willow Hight-Warburton (Centre for Inflammation Biology and Cancer Immunology, King's College London). Recombinant murine adipolin was gifted by Professor Noriyuki Ouchi (Department of Molecular Medicine and Cardiology, Nagoya University). The UK Biobank work (Approved Research ID 15147) was performed in collaboration with Dr Nick Dand (Department of Medical and Molecular Genetics, King's College London). Clinical code mapping to extract phenotypes was performed using software developed by Dr Alasdair Warwick (Institute of Cardiovascular Science, University College London). Calculation of genetic principal components, meta-analysis of the genome wide association studies and Bayesian fine-mapping of putative variants were performed by Dr Nick Dand and Dr Jake Saklatvala (Department of Medical and Molecular Genetics, King's College London).

Other support: I am extremely grateful to my PhD supervisors Professor John McGrath, Professor Maddy Parsons (Randall Centre for Cell and Molecular Biophysics, King's College London) and Dr Tanya Shaw for their guidance throughout my project. Tanya, I am ever so grateful for your boundless enthusiasm. I would

also like to thank Dr Nick Dand, Dr Joanna Jacków (St John's Institute of Dermatology, King's College London), Shaw Lab members (in particular Dr Willow Hight-Warburton, Deandra Belo De Freitas, Elena Drudi, Dr Fiona Kenny and Amy Lock), McGrath Lab members (in particular Dr Alexandros Onoufriadis, Stavroula Tekkela and Adam Sheriff), Parsons Lab members (in particular Sofia Endzhievskaya) and Dr Christina Philippeos (Centre for Stem Cells and Regenerative Medicine, King's College London).

Finally, a special thank you to Alasdair, Emily and my family - you keep me going!

Chuin Ying Ung
King's College London
30 September 2022

Abstract

Excessive scarring (keloids and hypertrophic scars) is complex with incompletely understood genetic underpinnings. The manifestation of a severe excessive scarring phenotype in an Israeli pedigree carrying a very rare homozygous frameshift variant (p.Asn43fs/rs571313759) in the *C1QTNF12* gene, encoding adipolin, provided an opportunity to study its potential role as a highly penetrant causal variant.

Both *in vitro* and bioinformatics approaches were adopted, with the latter extended to examine risk patterns of excessive scarring in a wider population. To characterise the phenotype of dermal fibroblasts carrying the variant, cells isolated from the proband were analysed for their morphology, proliferation, and contractility. These cells had enhanced transforming growth factor beta 1 (TGFβ1)-induced contractility but no differences in morphology or proliferation. To assess the effect of adipolin on fibroblast contractility, wild-type normal human dermal fibroblasts were treated with recombinant adipolin and contraction assays performed. No effect on intrinsic or TGFβ1-induced contractility was demonstrated.

To generate insight into the effect of adipolin on the fibroblast transcriptome, normal human dermal fibroblasts were treated with adipolin-conditioned media and bulk RNA-sequencing was performed. No significant transcriptomic changes were induced, suggesting that adipolin may not have a striking direct effect on dermal fibroblasts. To further substantiate the evidence for or against *C1QTNF12* loss-of-function (LOF) in fibrosis, exome data for ~200000 participants in the UK Biobank (UKB) was interrogated for potentially relevant phenotypic traits. No phenotypes of clinical relevance to excessive scarring were observed in individuals with *C1QTNF12* LOF variants.

Multivariable logistic regressions and a phenome-wide association study (PheWAS) was performed within the UKB excessive scarring cohort to characterise comorbid associations with excessive scarring. Previously reported associations with hypertension, vitamin D deficiency and atopic eczema were identified in models adjusting for age, sex and ethnicity, but only the association with atopic eczema (OR 1.68, $p<0.001$) was statistically significant after accounting for additional potential confounders. Ethnic differences in these comorbid associations were highlighted. The PheWAS identified numerous previously unreported disease associations that

may inform on common pathological mechanisms, including musculoskeletal disease and pain symptoms.

To investigate for common genetic variations that associated with excessive scarring, a genome-wide association study (GWAS) was performed for the UKB cohort and meta-analysed with GWAS summary statistics from FinnGen, comprising an independent European population. Three of four previously associated loci were identified: two at genome-wide significance (1q32.1: rs35383942, odds ratio [OR] 1.46, p-value 7.0×10^{-11}; 1q41: rs10863683, OR 0.69, p-value 3.0×10^{-25}) and one just short thereof (15q21.3: rs60890210, OR 1.24, p-value 6.7×10^{-8}). Bayesian fine-mapping offered strong evidence for causality of the lead SNP at 1q32.1 (posterior probability [PP] 0.79) and 1q41 (PP 1.0) but was inconclusive at 15q21.3. Expression quantitative trait loci (eQTL) were identified for ten genes including *PHLDA3* and *NEDD4* but colocalisation analysis did not support the hypothesis that these eQTLs shared a causal variant with keloid susceptibility. A known association with Dupuytren's contracture was identified at the 15q21.3 locus, notable as a fibrotic disorder, hinting at a shared genetic architecture.

Data from this work suggest that the *C1QTNF12* variant may not be a highly penetrant cause of excessive scarring in our pedigree but does not fully rule out a contributory role to the development of this complex disease. The UKB analyses were a broad examination of the phenotypic and genetic characteristics of excessive scars, with findings that parallel previous reports and shed light on new associations. Importantly, they highlighted the need for better representation of skin phenotypes in electronic health records for future research.

Contents

List of Figures

List of Tables

List of Abbreviations

2-D, 3-D . . . . . Two- or three-dimensional

3.1CM . . . . . pcDNA3.1-tagged-conditioned media (control conditioned media)

ABF . . . . . Approximate Bayes Factor

AdipCM . . . . . Adipolin-conditioned media

AKT . . . . . Protein kinase B

αSMA . . . . . Alpha-smooth muscle actin

BMI . . . . . Body mass index

BSA . . . . . Bovine serium albumin

CADD . . . . . Combined Annotation Dependent Depletion (tool)

CDM . . . . . Cell-derived matrix

cDNA . . . . . complementary DNA

CI . . . . . Confidence interval

COL1A1,COL7A1 Types 1 or 7 collagen

CTGF . . . . . Connective tissue growth factor

DANN . . . . . Deleterious Annotation of genetic variants using Neural Networks (tool)

DAPI . . . . . 4', 6-diamidino-2-phenylindole

DMEM . . . . . Dulbecco's Modified Eagle Medium

DNA . . . . . Deoxyribonucleic acid

dNTP . . . . . Deoxyribonucleoside triphosphate

EAF . . . . . Effect allele frequency

ECM . . . . . Extracellular matrix

EDTA . . . . . Ethylenediaminetetraacetic acid

EHR . . . . . Electronic health records

eQTL . . . . . Expression quantitative trait locus

ERK . . . . . Extracellular signal-regulated kinase

EVS . . . . . Exome Variant Server

ExAC . . . . . Exome Aggregation Consortium

FANTOM . . . . Functional ANnoTation Of the Mammalian genome (consortium)

FBS . . . . . Fetal bovine serum

FDR . . . . . False discovery rate

FG . . . . . FinnGen

FLAG . . . . . Anti-DYKDDDDK polypeptide

FUMA . . . . . Functional Mapping and Annotation (platform)

GFPCM . . . . . Green fluorescence protein-tagged-conditioned media (control conditioned media)

GRCh37/38 . . Genome Reference Consortium Human Build 37/38

GTEx . . . . . Genotype-Tissue Expression project

GWAS . . . . . Genome-wide association study

HLA . . . . . human leukocyte antigen

HPA . . . . . Human Protein Atlas

ICD . . . . . International Statistical Classification of Diseases and Related Health Problems

IGF . . . . . Insulin-like growth factor

IL1, IL6, IL10 . Interleukins-1 or 6 or 10

kDa . . . . . Kilodalton

KDF . . . . . Keloid fibroblasts

KLF . . . . . Krüppel-like factor

LB . . . . . Luria-Bertani

LD . . . . . Linkage disequilibrium

lincRNA . . . . . Long intervening noncoding RNA

LOF . . . . . Loss of function

MAF . . . . . Minor allele frequency

miR . . . . . MicroRNA

mRNA . . . . . Messenger RNA

NDF . . . . . Normal dermal fibroblasts

NF-κB . . . . . Nuclear Factor Kappa B

OR . . . . . Odds ratio

PBS . . . . . Phosphate buffered saline

PC . . . . . Principal component

PCOS . . . . . Polycystic ovarian syndrome

PCR . . . . . Polymerase chain reaction

PEI . . . . . Polyethylenimine

PheWAS . . . . Phenome-wide association study

PI3K . . . . . Phosphoinositide 3-kinase

PolyPhen-2 . . . Polymorphism Phenotyping v2 (tool)

PP . . . . . Posterior probability

QC . . . . . Quality control

QD . . . . . Quality by depth

qPCR . . . . . Quantitative polymerase chain reaction

QQ . . . . . Quantile-quantile

RNA . . . . . Ribonucleic acid

rpm . . . . . Revolutions per minute

scRNAseq . . . . Single-cell RNA sequencing

SIFT . . . . . Sorting Intolerant From Tolerant (tool)

SNP . . . . . Single nucleotide polymorphisms

TGFβ . . . . . Transforming growth factor beta

TNF . . . . . Tumour necrosis factor

UKB . . . . . UK Biobank

WES . . . . . Whole exome sequencing

WT . . . . . Wild type

1 Introduction

Contents

This chapter starts by reviewing the literature regarding the genetic and cellular abnormalities associated with excessive scarring (keloids and hypertrophic scars), with a focus on keloids. I then introduce a relatively unknown adipokine, adipolin, whose predicted homozygous loss-of-function (LOF) in a single pedigree with autosomal recessive excessive scarring, initiated my study of this disease at a cellular and population level.

1.1 Cutaneous scarring

Skin scarring is a natural part of wound healing, a tightly-regulated process comprising overlapping phases of inflammation, proliferation and remodelling(Marshall et al., 2018). After injury, a platelet and fibrin clot forms, inducing a complex inflammatory response, comprising numerous cellular and chemical mediators(Nurden, 2008). Neutrophils, typically the first cells to arrive, act by phagocytosis and cytokine release. They are followed by other immune cells, primarily macrophages and T-lymphocytes, carrying out reactions that overall serve to promote re-epithelialisation, remove debris and prepare the wound for the next phases of repair: proliferation and remodelling(Eming et al., 2007). This second phase serves to construct granulation tissue, contract the wound and reorganise the dermal extracellular matrix (ECM), and is orchestrated by fibroblasts, the most abundant cells in the dermis(Marshall et al., 2018). Granulation tissue, composed of ECM components such as collagens, proteoglycans, hyaluronic acid and elastin, as well as cytokines, growth factors and angiogenic factors, fills the physical defect caused by the injury, promotes further wound re-epithelialisation, and stimulates wound contraction(Marshall et al., 2018). A relatively acellular mass of disorganised ECM, devoid of dermal appendages (e.g. hair follicles, sweat glands) results(Marshall et al., 2018). At this point, wound resolution and tissue remodelling ensue to restore dermal strength and integrity. Inflammatory cell clearance and fibroblast apoptosis take place, accompanying a delicate balance of ECM synthesis and degradation(Shaw and Martin, 2009). A key phase by which scar outcome is determined, wound remodelling can last for over a year, although fully normal architecture and strength is never regained(Marshall et al., 2018).

While fetal skin is capable of scarless healing(Lorenz et al., 1992; Samuels and Tan, 1999), post-natal skin heals with scars — functionally-deficient tissue comprising denser, altered ECM with distortion of normal architecture(Xue and Jackson, 2015). A mature scar consists of a large amount of collagen with the majority being type 1 collagen and the rest being type 3 collagen(Monaco and

Lawrence, 2003). The collagen in normal skin is arranged in a non-parallel "basket-weave" orientation, whereas scar collagen is arranged in bundles parallel to the skin surface(Monaco and Lawrence, 2003). Scar tissue tends to be raised above the skin surface and hyperpigmented — these features tend to reduce as the scar matures(Baum and Arpey, 2006).

1.2 Excessive scarring

In susceptible individuals, imperfect regulation of wound resolution and remodelling results in excessive scarring(Wynn, 2008), encompassing keloids and hypertrophic scars. Hypertrophic scars are characteristically confined to the site of injury and may regress spontaneously, whereas keloids exhibit tumour-like properties, growing beyond the boundries of injury without regression and having a tendency to recur after excision(Limandjaja et al., 2020a). Whether they are distinct entities or a quantatitive extreme of normal wound scars remains an active area of debate. Histopathological distinction can be challenging(Gulamhuseinwala et al., 2008). A comprehensive review found that the most discerning histopathological features of keloids are a thickened, flattened epidermis; a tongue-like advancing edge in the dermis; haphazard, thick 'keloidal' collagen bundles; increased dermal cellularity; signs of inflammation; and variable alpha-smooth muscle actin (αSMA) expression(Jumper et al., 2015). However, most of these features can also be found in hypertrophic scars (Figure 1.1)(Limandjaja et al., 2020a). Similar abnormal and variable findings in ECM composition (e.g. increased levels of types 1 and 3 collagen, increased fibronectin, glycosaminoglycans, chondroitin sulfate and periostin, and decreased elastin), recently comprehensively reviewed(Limandjaja et al., 2020b), are reported for both excessive scar types. There is therefore a rationale for considering both scars as part of one entitiy. Accordingly, a recently introduced term, "keloid disorder", collectively refers to both scar types as well as keloidalis nuchae and acne keloidalis, conditions that also manifest with skin nodules as a result of trauma and inflammation(Uitto and Tirgan, 2020).

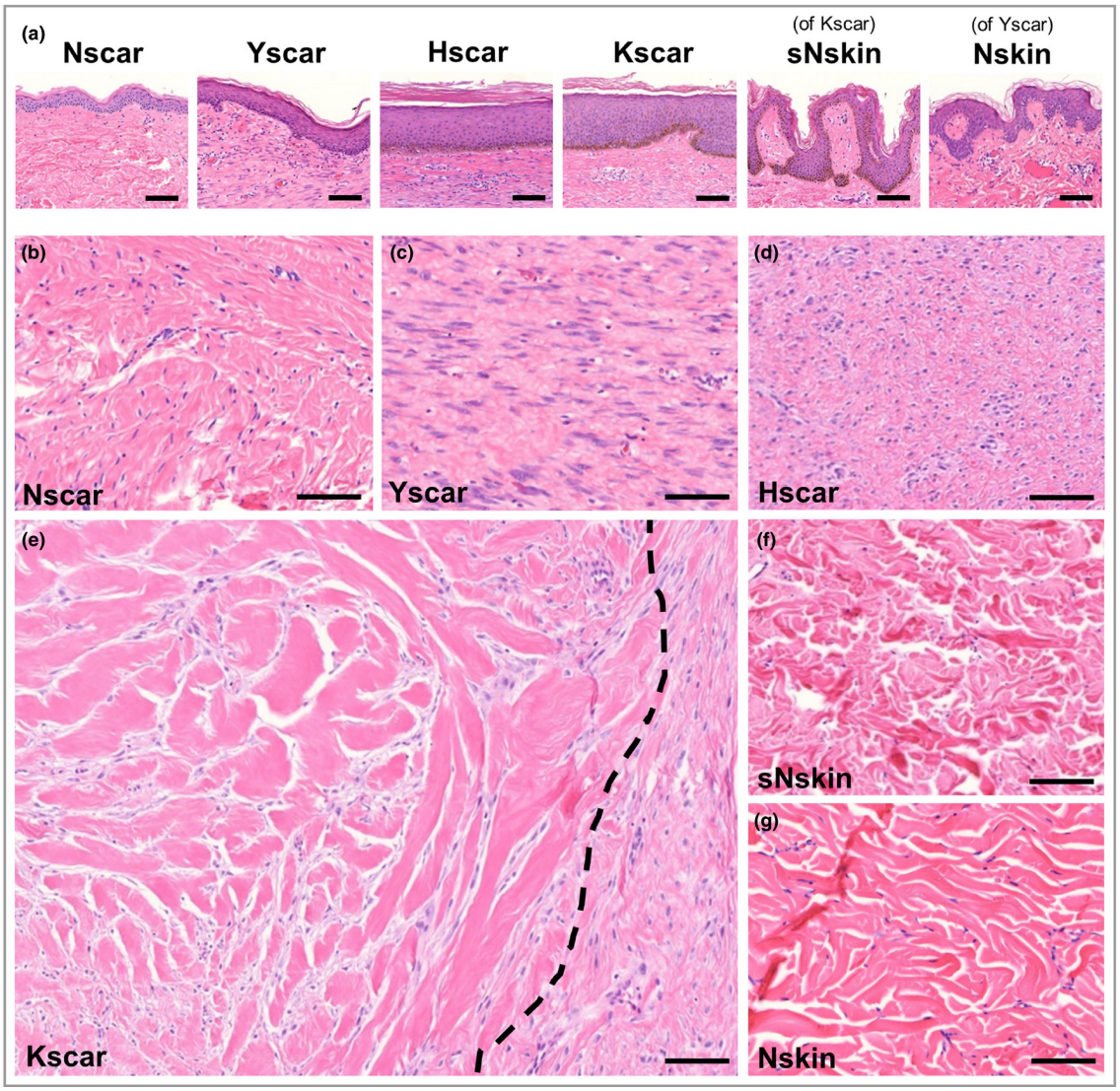

Figure 1.1: Histological characteristics of scars by representative haematoxylin and eosin staining. (a) Epidermis. (b–g) Dermis of (b) normal scar (Nscar), (c) young normal immature scar (Yscar) with (g) adjacent normal skin (Nskin), (d) hypertrophic scar (Hscar), and (e) keloid scar (Kscar) with (f) surrounding normal skin (sNskin). Scale bar = 100μm. The dashed line in (e) indicates the border between 'keloidal collagen' (left) and normal scar dermis (right). Open access figure reproduced from Limandjaja et al 2020.

These excessive scars provide a stark visual representation of fibrosis/fibro-proliferation, a process that inevitably besets most chronic diseases. Understanding their occurrence is of wide relevance. Numerous observations have been made, especially for keloids, each offering a window of insight to disease mechanism; however a unified understanding remains distant.

1.2.1 Keloid risk factors

Keloid risk factors can be grouped into patient-extrinsic and -intrinsic factors.

Patient-extrinsic factors

Patient-extrinsic factors relate to the type of injury that occurs (inciting stimulus) and the body site (topography) it affects. These factors increase the risk of excessive scar formation, regardless of individual susceptibility, and likely explain why not all susceptible individuals scar excessively after every injury(Slemp and Kirschner, 2006). Types of injuries that tend to lead to excessive scars are those that reach the dermis and/or cause its inflammation (e.g. trauma, piercings, insect bites, acne, burns)(Ogawa, 2017). This observation has led to numerous studies on the functional capacities of the heterogeneous resident cells in the context of scar formation; for example, studies on fibroblasts/fibroblast heterogeneity are burgeoning(Driskell et al., 2013; Rinkevich et al., 2015). Body sites that tend to form excessive scars include earlobes, neck, sternum, upper back, shoulders and upper limbs. Proposed reasons for why this may be the case include differences in skin tension(Bux and Madaree, 2010; Ogawa et al., 2012; Peacock et al., 1970), sebaceous gland density(Fong and Bay, 2002) and cellular embryological origins(Burd and Huang, 2005; Murray et al., 1981; Thulabandu et al., 2017).

Patient-intrinsic factors

Patient-intrinsic factors that influence excessive scar susceptibility have been uncovered largely through epidemiological observations. These broadly relate to the individual's ethnicity, genetics, gender, age, and comorbidities.

Ethnic differences Often-cited data on keloid incidence and prevalence is based largely on decades-old research, whereby ethnic differences are associated with skin tone(Bloom, 1956). In the black and Hispanic populations, incidence ranges from 4.5-16%(Bloom, 1956; Oluwasanmi, 1974; Rockwell et al., 1989), whereas in Taiwanese Chinese and whites, incidence is quoted at <1%(Bloom, 1956; Sun et al., 2014). However, observations that some lighter-skinned members of African American families can develop more severe keloids than their darker-skinned relatives(Marneros et al., 2001) and that keloids in Africans with albinism have a similar prevalence to the general population(Kiprono et al., 2015) suggest that increased pigmentation itself is not the sole explanation for the ethnic segregation of disease.

Genetics Although most keloids are sporadic, there exist numerous cases of familial keloids spanning black, white and Asian ethnicities, with observations of an increased prevalence in twins(Marneros et al., 2001; Ramakrishnan et al., 1974) and similarities in phenotypic presentation within families(Bayat et al., 2005a; Clark et al., 2009). The most commonly reported inheritance pattern is autosomal dominant with incomplete penetrance and variable expression(Chen et al., 2006a; Clark et al., 2009; Marneros et al., 2001; Ramakrishnan et al., 1974), although there have also been reports of autosomal recessive and X-linked inheritances(Goeminne, 1968; Omo-Dare, 1975). The most well-known genetic syndrome associated with keloids is Rubenstein-Taybi syndrome, caused by autosomal dominant mutations in the transcriptional co-activators *CREBBP* (chromosome 16p13.3) or *EP300* (chromosome 22p13) and is characterised by short stature, intellectual impairment, facial dysmorphism, broad thumbs and toes, and a higher incidence of keloids(Goodfellow et al., 1980; Kurwa, 1979; Selmanowitz, 1981; Shilpashree et al., 2015). It has been proposed that the association with keloids is through altered transcriptional co-activation of the TGFβ signalling pathway or altered expression of heat shock proteins(Lacombe and Morice-Picard, 2014). Interestingly, both *CREBBP* and *EP300* have also been shown to regulate profibrotic gene expression and the contractile phenotype in Dupuytrens disease, a fibrotic disease of the palmar fascia(Williams

et al., 2020). Other syndromes that have been reported to have keloids as a characteristic include the Ehlers Danlos syndromes (classical, linked to mutations in *COL5A1* and *COL5A2*(Burk et al., 2007) and type IV/vascular, linked to mutations in *COL3A1*(Ritelli et al., 2020)), the X-linked Goeminne syndrome(Zuffardi and Fraccaro, 1982) and a recently reported X-linked syndrome linked to *FLNA* mutations that manifests with joint contractures, keloids, large optic cup-to-disc ratios, and renal stones(Lah et al., 2016).

Age and gender Keloids are more frequently reported in women and in individuals between the ages of 10-30 years(Lu et al., 2014; Seifert and Mrowietz, 2009; Young et al., 2014). Although there is an argument that differences in health-seeking behavior may be contributory(Bayat et al., 2005a; Burd and Huang, 2005), observations of robust keloid growth in puberty and pregnancy(Ibrahim et al., 2020), periods of life during which sex-related hormones spike, hint at an endocrinological contribution to disease susceptibility(Glass, 2017). However, available literature is scant. Oestrogenic and gonadotrophic substances have been observed in keloid tissue and shown to stimulate connective tissue growth(Geschickter and Lewis, 1935), potentially through the stimulation of TGFβ1 signalling(Son et al., 2005). Sex hormones may also contribute to increased susceptibility through their effect on immune regulation. For example, oestrogen has been reported to be pro-inflammatory by stimulating the AKT/mTOR pathway(Calippe et al., 2010; Pratap et al., 2015), which is often implicated in keloidogenesis(Syed et al., 2012).

Comorbidities There are numerous anecdotal reports and independent studies reporting diseases that tend to co-exist in individuals with excessive scarring. To date, these have been based on epidemiological observations and speculated biological mechanisms. The most prominently discussed disease associations include hypertension(Rutherford and Glass, 2017; Woolery-Lloyd and Berman, 2002), uterine leiomyoma(Sun et al., 2014), atopic eczema(Kwon et al., 2021; Lu et al., 2018), vitamin D deficiency(El Hadidi et al., 2021; Yu et al., 2013) and skin cancer(Y.-Y. Lu

et al., 2021). A comprehensive, unbiased examination of the systemic comorbidities of individuals susceptible to excessive scarring, which may support the discovery of new mechanistic relationships, is lacking.

1.2.2 Keloid genetics

The patient-intrinsic factors described above, namely, ethnic differences, familial clustering and genetic syndromes, suggest a genetic basis to keloid pathogenesis. Familial and population studies have identified several loci to be associated with keloids. It is still unclear whether these loci are ethnicity specific or whether there are common elements in the genetics of keloids across ethnicities. The mechanisms underlying these associations remain unknown and it is likely that there is a substantial degree of genetic heterogeneity underlying keloids. Key genetic studies performed to date are summarised.

Linkage studies

Several susceptibility loci have been identified through familial linkage studies, including chromosome 2q23 (suggested gene *TNFAIP6*) unique to a Japanese family and 7p11 (suggested gene *EGFR)* in an African American family(Marneros et al., 2004). Associations with these loci were not reproduced however, when other keloid pedigrees were screened(Chen et al., 2006b; Marneros et al., 2004). Indeed, other susceptibility loci (on chromosome 10q21.21 and 18q21, implicating *SMAD2, SMAD4* and *PIAS2* respectively*)* were identified in Chinese Han pedigrees(Chen et al., 2006b; Yan et al., 2007).

Candidate gene association analyses

Key candidate genes studied are related to TGFβ, *TP53* and the human leukocyte antigen (HLA).

TGFβ Despite differences in the expression of TGFβ genes between keloid and normal fibroblasts (higher expression of *TGFB1*, lower expression of *TGFB3* in keloids)(Bock et al., 2005), studies performed, primarily from one research group, have found no associations between keloid and polymorphisms or mutations in these genes or their downstream signalling molecules (*SMAD3*, *SMAD6* and *SMAD7*)(Bayat et al., 2005b, 2002; Bayat et al., 2004, 2003; Brown et al., 2008a). However, this does not exclude the possibility that there are relevant polymorphisms affecting upstream genes and/or long-range modulators of the TGFβ pathway(Glass, 2017).

TP53 Being important in cell proliferation and apoptosis, this gene has been proposed to influence keloid predisposition. As with TGFβ, although lesional expression differences have been demonstrated(Ladin et al., 1998), no convincing genetic associations have thus far been found. A meta-analysis of six Chinese studies (359 keloid cases and 493 healthy controls) showed no association between the commonly studied *TP53* polymorphism (rs1042522) and keloids(Wu et al., 2012). A study of a Polish cohort (86 cases, 100 controls) similarly found no association between *TP53* polymorphisms (rs1042522 or rs17878362) and keloid risk(Dmytrzak et al., 2019).

HLA The HLA system has been associated with other dermal fibrotic diseases e.g. sarcoidosis(Grunewald et al., 2010). Two groups have studied the association between *HLA-DRB1* alleles and keloids, in white (67 cases, 537 controls)(Brown et al., 2008b) and Chinese (192 cases, 273 controls)(Lu et al., 2010) respectively, both supporting an association with *HLA-DRB1*15*. This suggests an immunogenic component to keloids, although this finding has not been replicated in population-wide genome scans.

Genome-wide population studies Three genome-wide association studies (GWAS) in Japanese (816 cases, 2385 controls)(Ishigaki et al., 2020; Nakashima et al., 2010) and Chinese Han (714 cases, 2944 controls)(Zhu et al., 2013) populations have independently identified two shared loci to be associated with keloids, namely 1q41(Ishigaki et al., 2020; Nakashima et al., 2010; Zhu et al., 2013) and 15q21.3(Ishigaki et

al., 2020; Nakashima et al., 2010; Zhu et al., 2013). The Japanese studies additionally identified associations on chromosome 3 which were not identified in the Chinese study. A cross-population GWAS meta-analysis combined keloid cases from Japanese(Nagai et al., 2017) (1055 cases, 177671 controls) and European populations(Kurki et al., 2022; Sudlow et al., 2015) (668 cases, 481244 controls) and identified 2q37.3 as a novel risk locus (although no significant association was identified in the individual studies analysed)(Sakaue et al., 2021).

Only the single nucleotide polymorphisms (SNPs) in the 15q21.3 and 3q22.3–23 loci were within genes: rs8032158 in 15q21.3 maps to *NEDD4* (an E3 ubiquitin ligase which has been implicated in TGFβ and IGF1 signalling)(Nakashima et al., 2010) whereas rs1511412 in 3q22.3-23 maps to *FOXL2* (a transcription factor that stimulates the expression of gonadotrophin-releasing hormone and regulates cholesterol metabolism and steroidogenesis)(Nakashima et al., 2010).

Admixture mapping association analysis in an African American cohort (122 cases, 356 controls)(Velez Edwards et al., 2014) showed an association on chromosome 15q21.2-22.3, which includes *NEDD4*, implicated in both Asian GWASs, although the most significantly associated SNP (rs747722) was within a different gene, *MYO1E* (a non-muscle myosin involved in actin crosslinking). Associations were also found for SNPs at 11q13.5 (rs35641839 in *MYO7A* another non-muscle myosin) and 4q31.21(rs2636675)(Velez Edwards et al., 2014). However it should be noted that none of the associations in this study met Bonferroni significance.

1.2.3 Cellular keloid abnormalities

The predisposing factors described above culminate in cellular abnormalities characteristic of and ultimately responsible for keloid predisposition. Diverse cell types have been implicated. For example, keloid keratinocytes have been shown to contribute paracrine regulation of fibroblasts through wound healing mediator secretion(Khoo et al., 2006; Lim et al., 2002; Ong et al., 2006) and undergo epithelial-mesenchymal transition(Marconi et al., 2021; Stone et al., 2016; Yuan et al., 2019). Endothelial cell dysfunction has been suggested to play a role in keloidogenesis

through increased vascular permeability (prolonging the influx of inflammatory mediators)(Ogawa and Akaishi, 2016) as well as endothelial-mesenchymal transition, serving as a source of abnormal keloid fibroblasts(Lee et al., 2016). Increased levels of immune cells including macrophages (with anti-inflammatory pro-fibrotic features)(Bagabir et al., 2012a; Boyce et al., 2001; Jin et al., 2018), T-lymphocytes(Bagabir et al., 2012a; Boyce et al., 2001) and mast cells(Ammendola et al., 2012; Bagabir et al., 2012a) have been reported in keloids. A common theme underlying these various cellular abnormalities in keloids is their effect on the dermal fibroblasts. These cells are reviewed further in the context of keloids/excessive scarring(Wang et al., 2020).

The keloid fibroblast

The keloid dermal fibroblast (KDF), as the main ECM-producing dermal cell type, has been the overwhelming focus of keloid research. Although frequently described as myofibroblasts (αSMA-expressing contractile fibroblasts that overproduce ECM), these cells are not necessarily universally present in keloids and there are numerous inconsistencies on the features/identity of KDFs(Bell and Shaw, 2021). Differentiating features of KDFs that are most often reported in the literature are their proliferation, contractility, ECM synthesis, and cytokine expression(Chipev and Simon, 2002; Limandjaja et al., 2020b; Mukhopadhyay et al., 2005).

A general consensus from most *in vitro* monolayer studies is that KDFs are hyperproliferative(Blume-Peytavi et al., 1997; Ghazizadeh et al., 2007; Hanasono et al., 2003; Lim et al., 2006; Romero-Valdovinos et al., 2011; Russell et al., 1988; Xin et al., 2017) with reduced apoptosis. This is thought to contribute to a larger than normal pool of cells and the resultant excessive ECM(Chipev et al., 2000; Ladin et al., 1998; Luo et al., 2001; Messadi et al., 1999). However, a more complicated picture arises when considering intralesional heterogeneity. Research groups generally support the concept of quiescent (hypoproliferative) central keloid regions, with active (hyperproliferative) peripheries(Suttho et al., 2016; Varmeh et al., 2011); however, there exist others who view the keloid centre as a driver of pathology, demonstrating increased proliferation and a lack of apoptosis(Akasaka

et al., 2001; Giugliano et al., 2003; Tsujita-Kyutoku et al., 2005). The variability of these findings may reflect differences in defining lesional centre and peripheries as well as the heterogeneity of the keloid phenotype itself (e.g. 'superficial spreading' lesions may have quiescent centres whereas 'bulging' lesions may have aggressive cores)(Limandjaja et al., 2020b).

In line with the theory that keloid cores play an active role in disease, *in vitro* reconstructed keloid lesions (comprising patient-derived keratinocytes and fibroblasts) have demonstrated a more exaggerated phenotype in the central deep keloid region, in particular, with increased contraction(Limandjaja et al., 2018b). In this region, activated fibroblasts/myofibroblasts were found to be more abundant and thought to account for the increased contractility. The observation of a more contractile KDF phenotype has been supported by other studies, proposing that these fibroblasts are more primed to effectively transduce TGFβ1 signalling(Chipev and Simon, 2002; Kamamoto et al., 2003; Saito et al., 2011).

Numerous cytokines are implicated in keloids, chiefly TGFβ1(Bettinger et al., 1996; Bock et al., 2005; Daian et al., 2003; Fang et al., 2016; Fujiwara et al., 2005; Hanasono et al., 2003; Messadi, 1998; Mikulec et al., 2001; Xia et al., 2006, 2004), connective tissue growth factor (CTGF)(Fang et al., 2016; Xia et al., 2007, 2004), platelet-derived growth factor (PDGF)(Haisa et al., 1994), vascular endothelial growth factor(Fujiwara et al., 2005; Ong et al., 2006; Wu et al., 2004) and interleukin-6 (IL6)(Ghazizadeh et al., 2007; Xue et al., 2000). Interpreting their individual contributions to pathological fibroblast behaviours amidst the cross-regulatory networks is a growing challenge. The TGFβ family has been extensively studied in relation to fibrosis, with TGFβ1 and TGFβ2 being widely-accepted as profibrotic factors and TGFβ3 being anti-fibrotic(Shah et al., 1995). Both canonical (via the SMAD complex) and non-canonical (e.g. via ERK/MAPK) TGFβ1 signalling can drive the differentiation and persistence of myofibroblasts, which in addition to their increased contractility, can produce inappropriate ECM(Clayton et al., 2020). More recent mechanisms by which TGFβ1 has been shown to promote myofibroblast/KDF differentiation and activity include through the DNA, RNA and protein-regulatory

functions of non-coding RNA (e.g. microRNA, miR-21(Liu et al., 2016), long non-coding RNAs, HOXA11-AS(Jin et al., 2019) and GAS5(Tang et al., 2020)).

The abnormal cellular behaviours and molecular mechanisms of the KDF culminate in the synthesis of a qualitatively-distinct and excessive ECM. Whereas scarless fetal ECM is rich in type 3 collagen and hyaluronic acid(Hallock et al., 1988; Reinke and Sorg, 2012), adult scars are rich in type 1 collagen and this enrichment is exaggerated in keloids(Xue and Jackson, 2015). A discovery-based proteomics study further characterised the keloid ECM, finding specific enrichment in proteins (e.g. aggrecan, asporin, biglycan, cartilage-associated protein, collagen α-1 (XI) chain), that when functionally grouped, revealed changes in ECM assembly, a loss of degradative proteases and a cartilage-like phenotype(Barallobre-Barreiro et al., 2019). The chondrogenic potential of dermal fibroblasts is established(Outani et al., 2013) and KDFs have been reported to express chondrogenic genes(Naitoh et al., 2005). Whether these findings reflect their cellular origin is an area of ongoing research.

The study of fibroblast plasticity and heterogeneity is fast-evolving. Beyond myofibroblasts, fibroblast characterisation is being increasingly refined. Within the last two years, there have been three publications of single cell RNA-sequencing (scRNAseq) studies on keloids(Deng et al., 2021; Liu et al., 2022; Shim et al., 2022). Deng et al described four main fibroblast subpopulations that emerged from 13 clusters: secretory-papillary, secretory-reticular, pro-inflammatory and mesenchymal, the latter of which (characterised by high expression of periostin, cartilage oligomeric matrix protein, type 11 collagen and asporin), is most relevant in keloid pathgenesis(Deng et al., 2021). Shim et al and Liu et al both confirmed previously defined expression markers (e.g. periostin) of KDF subpopulations and independently highlighted the significance of fibroblast and vascular subpopulations as well as their associated dysregulation of TGFβ signalling(Liu et al., 2022; Shim et al., 2022).

The keloid-prone fibroblast

Although interest in the KDF is growing, there is a paucity of research related to 'keloid-prone' fibroblasts. No consensus exists on how they should be defined - whether they are represented by fibroblasts from the non-lesional (distant or peri-lesional/keloid-adjacent) skin of individuals with keloids or by fibroblasts from keloid-prone anatomical sites. As alluded to previously, a keloid-prone individual does not necessarily form keloids after every injury, yet not all individuals form keloids in keloid-prone anatomical sites(Al-Attar et al., 2006; Fong et al., 1999; Slemp and Kirschner, 2006).

Studies that have attempted to differentiate scar fibroblasts have performed three main comparisons: (i) fibroblasts from 'scar-prone' adult skin versus 'less scar-prone' neonatal skin(Mateu et al., 2016; Živicová et al., 2017); (ii) fibroblasts from keloid-adjacent skin versus keloid scars and normal skin (from keloid-unaffected individuals)(Ashcroft et al., 2013; Limandjaja et al., 2018a); and (iii) fibroblasts from 'scar-prone' truncal fibroblasts versus 'less scar-prone' facial fibroblasts(Kurita et al., 2012). Most findings align with observations of the KDF itself; for example, 'scar-prone' fibroblasts from adult skin are less plastic; 'scar-prone' keloid-adjacent fibroblasts have increased proliferation, contractility and show a similar keloid-ECM profile to KDFs(Ashcroft et al., 2013; Limandjaja et al., 2018a); 'scar-prone' truncal fibroblasts have a keloid-fibroblast cytokine profile (greater type 1 collagen, TGFβ1, CTGF expression)(Kurita et al., 2012). However, other findings such as the greater expression of αSMA by 'less scar-prone' neonatal fibroblasts(Mateu et al., 2016; Živicová et al., 2017) are counter-intuitive.

Overall, these findings offer hints for a keloid-prone fibroblast that induces keloid formation upon receipt of an inciting stimulus, but they are not direct surrogates. A clearer definition along with advances in cellular characterisation may aid the efforts into researching this.

1.2.4 Emerging/underexplored keloid hypotheses

The challenges in understanding keloids have led to a growing list of proposed pathomechanisms, some of which are discussed here.

Keloid epigenetics

Epigenetic modifications (alterations to DNA expression without affecting its sequence, including DNA methylation, histone modification and non-coding RNA regulation, that ultimately lead to control of gene expression and phenotype) have been reported in keloids, including altered DNA methylation in keloid tissue and fibroblasts(Jones et al., 2015; Russell et al., 2010) and differential expression of epigenetic modifying enzymes and non-coding RNAs (e.g. miRNA and lincRNA)(Fitzgerald O'Connor et al., 2012; Guo et al., 2016; Stevenson et al., 2021). DNA methylation can either suppress or activate gene expression (by inhibiting transcription factor or repressor complex binding) whereas non-coding RNA sequences regulate mRNA both transcriptionally and post-transcriptionally(Stevenson et al., 2021). Overall, the functional implications of the epigenetic changes observed in keloids are poorly understood. Of the studies performed, the resulting altered processes were associated with TGFβ signalling(Hu et al., 2017; Wu et al., 2019; Zhang et al., 2015), WNT signalling(Sun et al., 2017), fibroblast proliferation(Feng et al., 2017; Liu et al., 2014; Zhang et al., 2016) and ECM formation(Wu et al., 2019; Yuan et al., 2021).

Vasculature in keloids

Several lines of evidence suggest that vascular dysfunction may be relevant to the development of keloids. These include abnormal microvasculature in keloids, the enrichment of vascular subpopulations in keloid single-cell transcriptomics studies, the co-occurrence of keloids and hypertension(Arima et al., 2015, 2015; Dustan, 1995; Lawton Snyder et al., 1996; Noishiki et al., 2017), and the fact that many treatment options for keloids suppress angiogenesis (e.g. radiotherapy(Grabham and Sharma, 2013), laser therapy(Koike et al., 2014), compression therapy(Ogawa and Akaishi,

2016), corticosteroids(Marks et al., 1982)). This theory suggests endothelial dysfunction allows passage of inflammatory mediators and induces abnormal fibroblast function(Ogawa and Akaishi, 2016).

Inflammation in keloids

Keloids are known to exhibit chronic inflammation within their expanding margins; infiltrating inflammatory cells as well as an abundance of pro-inflammatory factors (e.g. IL6, cyclooxygenase-1) have been demonstrated(Abdou et al., 2014; Bagabir et al., 2012a; Huang et al., 2012; Jumper et al., 2015). Prolonged inflammation has been shown to increase fibroblast activities, resulting in greater and more sustained ECM deposition(Jumper et al., 2015). Consistent with these findings, anti-inflammatory molecules (e.g. IL10) have been shown to negatively regulate ECM synthesis and reduce scar formation(Kieran et al., 2013; Shi et al., 2016). Furthermore, corticosteroids, the most widely-used scar treatment, is a well-known anti-inflammatory agent(Amini-Nik et al., 2018).

However, this finding seems to conflict with the general understanding of the inflammatory processes in wound healing and tissue repair. In the early stages of wound healing, pro-inflammatory factors are required as defence mechanisms, whereas in later stages, anti-inflammatory factors are dominant, promoting cellular proliferation and remodeling(Wang et al., 2020). If cellular proliferation underlies excessive scar formation, one would assume that the sustained anti-inflammatory phase is pathogenic. Interestingly, one of the most-implicated immune cell types in keloids is the M2 macrophage, which is renowned for its anti-inflammatory functions but also highly relevant in tissue repair and remodeling — amongst the numerous anti-inflammatory cytokines it secretes is TGFβ1(Barros et al., 2013; Hesketh et al., 2017; Jin et al., 2018; Seoudy et al., 2022).

Metabolism in keloids

I have previously reviewed the metabolic aberrations found in fibrotic diseases(Ung et al., 2021). Briefly, the role of metabolism in keloids is highlighted through

studies on metabolic activity(Ueda et al., 1999), glucose and lipid metabolism(Louw, 2000; Tachi and Iwamori, 2008). It is also indirectly alluded to through discussions on the involvement of adipokines (primarily adiponectin) in fibrosis, including keloids(Darmawan et al., 2020; Luo et al., 2017; Żółkiewicz et al., 2019).

High-performance liquid chromatography of keloid scars detected increased fibroblast number, along with high levels of adenosine triphosphate(Ueda et al., 1999). Further exploration uncovered that the fibroblasts had enhanced glycolysis and attenuated oxidative phosphorylation, through altered expression of glucose uptake receptors and mitochondrial enzymes(Vinaik et al., 2020; Vincent et al., 2008). These changes may provide a mechanism for the cancer-like behaviour of KDFs(Vincent et al., 2008) and are thought to cause similar effects including altered signal transduction pathways (e.g. PI3K/AKT) which impact cellular development and growth(Wang et al., 2021).

Lipid profile analyses of keloids showed differences in their lipid composition including a lower proportion of triglycerides, cholesterol and wax esters(Tachi and Iwamori, 2008) and a higher proportion of arachidoic acid(Louw, 2000). This and the fact that corticosteroids can alter lipid synthesis(Kao et al., 2003) support their involvement in keloid development(Huang and Ogawa, 2013). Various hypotheses have been put forward as to their pathomechanism including the modulation of inflammation and signal transduction through lipid-derived mediators as well as the modulation of mechanotransduction, the process by which physical/tension forces are converted into biochemical signals that are intergrated into cellular responses, through lipid rafts(Huang and Ogawa, 2013).

A major group of secretory proteins that modulate cellular metabolism are adipokines, which are mainly produced by adipose tissue. There is an increasing list of adipokines that regulate numerous aspects of energy homeostasis; their imbalance has been associated with a range of biochemical alterations resulting in disease(Deng and Scherer, 2010). Most prominent amongst 'beneficial' adipokines is adiponectin, a pleiotropic adipokine that targets multiple organ systems to regulate insulin sensitivity, energy balance, and cellular metabolism(Wang and Scherer, 2016). Its

importance has even been extended to the regulation of inflammation, wound healing and fibrosis(Fang et al., 2012; Luo and Liu, 2016; Reinke et al., 2016; Shibata et al., 2012). Observations that keloids were found to express less adiponectin and its receptors were followed by findings that adiponectin suppressed CTGF-induced KDF proliferation, migration and ECM production(Luo et al., 2017). The adiponectin discoveries have prompted a search for other 'good adipokines' as therapeutic targets for numerous diseases, including for skin fibrosis(Żółkiewicz et al., 2019).

1.3 Adipolin

Adipolin is an adipokine first identified in 2011 by two independent laboratories(Enomoto et al., 2011; Wei et al., 2012b). In a discovery study for novel adipokines, Enomoto et al identified adipolin through a microarray analysis of gene expression profiles of epididymal adipose tissues in lean versus obese mice, finding it to be one of the transcripts downregulated in adipose tissues of obese mice. Independently, in searching for novel factors with sequence homology to the more well-known adipokine, adiponectin, Wei et al performed an *in silico* screen and identified, amongst other multimeric proteins, adipolin, which they designated as CTRP12.

1.3.1 Adipolin gene and protein structure

The gene encoding adipolin is most widely referred to as *C1QTNF12* (also referred to as *CTRP12*, *C1qdc2*, *FAM132A*) and this nomenclature will be used hereafter. Human *C1QTNF12* has been mapped to chromosome 1p36.

The protein-coding transcript for adipolin (ENST00000330388) has 1036 base pairs, translating to 302 amino acids, encoding eight coding exons. It is predicted to have a signal peptide, a cleavage site, and a complement component 1q/tumour necrosis factor (TNF)-like domain (Figure 1.2)(Cunningham et al., 2021).

The adipolin protein likely undergoes some post-translational modifications as suggested by its structure – human adipolin contains a glycosylation site at amino acid position 43 and a furin cleavage site at amino acid position 97-98; cleavage

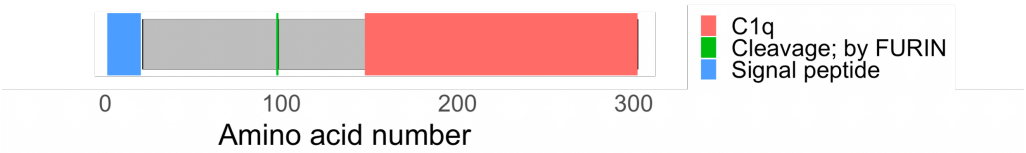

Figure 1.2: Key domains for the adipolin coding transcript. C1q, complement 1q. Image produced using the R package drawProteins based on data from the Uniprot Protein Database.

is predicted to produce a globular isoform of the full-length protein(Bateman et al., 2020). Proteolytic cleavage of adipolin is likely to be performed by the serine protease, furin, which is highly expressed in adipose tissue, and this process is enhanced by insulin(Wei et al., 2012a). In keeping with these *in silico* predictions, endogenous adipolin secreted from differentiated 3T3-L1 adipocytes, circulating adipolin in human and mouse serum as well as FLAG epitope-tagged adipolin expressed in mammalian HEK293T cells have been shown to exist in two isoforms: full-length (~40kDa) and globular (~25kDa)(Wei et al., 2012a). The globular isoform of adipolin appears to dominate in the circulation(Wei et al., 2012a).

Subunit interactions for human adipolin are inferred from its murine homolog. Through disulfide-linked structures, full-length adipolin is thought to form a higher molecular mass oligomer (~120kDa or higher) consisting predominantly of trimers and larger complexes, whereas globular adipolin forms a lower molecular mass oligomer (~45kDa) consisting predominantly of dimers(Wei et al., 2012a).

1.3.2 Adipolin expression

Tissue expression

The murine adipolin transcript is reported to be widely expressed, notably in adipose tissue(Enomoto et al., 2011; Wei et al., 2012b), kidneys(Enomoto et al., 2011; Wei et al., 2012b), intestines(Enomoto et al., 2011) and testes(Wei et al., 2012b).

In humans, discrepancy exists regarding tissue expression of the adipolin transcript. Using human tissue cDNA panels comprising 16 tissue types and 10 blood cell types, Wei et al found that adipolin is expressed chiefly by adipose tissue, over 50-fold greater than the next highest expressing tissue, the kidneys, and minimally

expressed in the other tissues in the panels(Wei et al., 2012b). In the Human Protein Atlas (HPA) Consensus dataset, a resource created by the combination of three transcriptomics datasets (HPA(Uhlén et al., 2015), GTEx and FANTOM5(Lizio et al., 2018)) for 55 tissue types and six blood cell types, the adipose tissue expression of adipolin was less marked, with transcript expression enhanced in the intestine and skin, followed by the kidneys, testes and granulocytes(Uhlén et al., 2015). This is in part supported by immunohistochemical staining, showing membranous and cytoplasmic positivity on colonic glandular cells, however no convincing immunohistochemical validation was seen in skin.

Cellular expression

Few independent studies have reported on cellular expression of adipolin. Its transcript expression has been detected in human primary chondrocytes, the T/C-28a2 (human chondrocyte) cell line and the ATDC-5 (murine chondrocyte) cell line(Conde et al., 2012). At a protein level, adipolin has been detected in 3T3L1 adipocyte media(Enomoto et al., 2012) and lysates(Wei et al., 2012a) as well as in H9c2 cardiomyocytes(Zhou et al., 2020), although in the latter study, antibody details were not provided.

Again, some discrepency arises when cellular expression of the adipolin transcript is examined using data from the HPA. Assessed at a single cell transcriptome level for 13 tissues (excluding adipose tissue) and peripheral blood mononucleated cells (using datasets from the Single Cell Expression Atlas(Madeira et al., 2022), Human Cell Atlas(Regev et al., 2017), Gene Expression Omnibus(Edgar, 2002) and European Genome-phenome Archive(Freeberg et al., 2021)), the adipolin transcript appears to be group-enriched in renal proximal tubular cells and adipocytes. Notably, using this single cell resource, there was no detectable expression in the colon or the skin, unlike the tissue level datasets(Karlsson et al., 2021). Furthermore, there is insufficient protein validation of adipolin expression, with western blotting failing to detect convincing expression in the standard panel of cell lysates tested(Uhlén et al., 2015).

1.3.3 Regulation of adipolin biosynthesis/expression

In healthy humans, plasma adipolin levels have been measured in the range of 749 - 1,176pg/ml(Babapour et al., 2020; Bai et al., 2017; Fadaei et al., 2019; Tan et al., 2014a). This is substantially lower than the more well-known adipokines, adiponectin (~10μg/ml(Gavrila et al., 2003)) and leptin (~17 - 38ng/ml(Hellstrom et al., 2000)) but higher than ghrelin (~357 - 522fg/ml)(Tschop̈ et al., 2001).

There are conflicting reports regarding sex-specific adipolin expression; an initial study in healthy subjects reported similar expression levels (~700pg/ml – ~1150pg/ml, $p>0.05$)(Tan et al., 2014b) but a more recent study in patients with coronary artery disease and healthy controls reported higher levels in females compared to males when covariate effects were removed (761 $\pm$ 253pg/ml versus 596 $\pm$ 213pg/ml, $p<0.001$)(Fadaei et al., 2019). Collectively, most published studies have reported lower adipolin levels in cardiometabolic diseases, namely type 2 diabetes, cardiovascular disease and polycystic ovarian syndrome (PCOS). In patients with cardiovascular disease, adipolin levels have been measured in the range of 609 - 929pg/ml(Babapour et al., 2020; Fadaei et al., 2019) with levels particularly decreased in patients with acute myocardial infarction and stable angina pectoris.(Babapour et al., 2020) In patients with PCOS, levels were lower, at 338pg/ml.(Tan et al., 2014a) For type 2 diabetes however, study results are inconsistent. Bai et al(Bai et al., 2017) reported significantly lower levels (387.7 - 506.4pg/ml) compared with healthy controls (753.8 - 892.4pg/ml). However, Kasabri et al(Kasabri et al., 2019) found significantly higher levels in patients with normoglycaemic (1.08 $\pm$ 0.96 ng/ml) and pre-diabetic (1.41 $\pm$ 0.77ng/ml) metabolic syndrome. In support of the latter study, Mehrdadi et al(Mehrdadi et al., 2016) found that coenzyme Q supplementation reduced hemoglobin A1c levels in overweight patients with type 2 diabetes while simultaneously reducing adipolin levels. Moreoever, a further study comparing adipolin levels in patients with type 2 diabetes and patients with and without type 2 diabetes undergoing haemodialysis, found significantly higher levels in the patients with diabetes (29ng/ml)(Alipoor et al., 2019).

As alluded to above, adipolin expression seems to be modulated by metabolic alterations, although there are also disparate findings between murine and human studies. Its transcriptional expression is significantly reduced in the adipose tissue of diet-induced obese mice(Enomoto et al., 2011; Wei et al., 2012b) and although both globular and full-length adipolin are reduced in obesity, there was an increase in the ratio of globular to full-length adipolin, indicating enhanced cleavage in diet-induced obese mice(Enomoto et al., 2012). In accordance with the notion that obesity is associated with low grade inflammation, treatment of cultured adipocytes with the pro-inflammatory cytokine TNF-α led to reduced adipolin expression(Enomoto et al., 2011). Contrastingly, analyses of the adipose tissue of otherwise healthy obese women showed they had increased adipolin transcript expression and that this correlated positively with the expression of TNF-α(Omidifar et al., 2019). With fundamental differences between the murine and human studies, including distribution and levels of adipolin expression as well as chronicity of obesity in the subjects studied, it is difficult to make direct cross-references. One may postulate that in the early phases of obesity, adipolin expression is suppressed, and over time its expression is upregulated as a protective measure against insulin resistance(Omidifar et al., 2019).

Pharmacological interventions that alter insulin sensitivity have been shown to affect adipolin levels in mice and humans. Wei et al(Wei et al., 2012b) first reported that insulin and the insulin-sensitising agent rosiglitazone increase adipolin transcript and protein expression in adipose tissue. This finding was recapitulated in human studies by Tan et al(Tan et al., 2014a; Tan et al., 2014b). Hyperinsulinaemia induction in healthy subjects with an insulin-glucose infusion led to a marked increase in adipolin levels from ~1ng/ml to ~1.5 - 1.9ng/ml for the duration of the infusion (26 hours) whilst blood glucose levels remained steady(Tan et al., 2014b). The authors also showed that the addition of insulin to control human subcutaneous adipose tissue explants led to increased adipolin protein expression and secretion, and that these effects were significantly attenuated by the PI3K inhibitor, LY294002, but not the MEK inhibitor, U0126. They then demonstrated a similar stimulatory

effect on adipolin expression using the insulin-sensitising peroxisome proliferator-activated receptor (PPAR)γ agonist, rosiglitazone, and blocked this effect using the PPARγ inhibitor GW9662 (10μM). In a separate study, treating women with PCOS with a different insulin-sensitiser, metformin, for six months increased their adipolin levels from 388.7pg/ml to 922.2pg/ml(Tan et al., 2014a). In contrast, oral glucose challenges reduced circulating adipolin levels (from 447pg/ml to 282pg/ml (P<0.01) in patients with type 2 diabetes(Bai et al., 2017), from ~550pg/ml to ~500pg/ml in patients with PCOS(Tan et al., 2014a) and from 447pg/ml to 282pg/ml in healthy controls(Tan et al., 2014a)).

Little is known about the transcriptional regulation of adipolin. To date, two members of the Krüppel-like factor (KLF) family of zinc-finger transcription factors have been implicated. Bell-Anderson et al(Bell-Anderson et al., 2013) identified *KLF3* as a transcriptional repressor of murine adipolin (*C1qtnf12*) whereas Enomoto et al(Enomoto et al., 2013) identified *KLF15* as its transcriptional enhancer. Consistent with a metabolic phenotype affecting adipolin expression, *KLF3*-null mice are lean and protected from diet-induced obesity(Bell-Anderson et al., 2013), whereas *KLF15* expression is decreased in adipose tissue of obese mice and TNFα treated adipocytes(Enomoto et al., 2013). However, mouse and human genomes have a substantial divergance of sequences involved in transcriptional regulation(Yue et al., 2014); although the mouse *C1qtnf12* promoter region contains several *KLF3*-binding sites(Bell-Anderson et al., 2013), examination of the human *C1QTNF12* proximal promoter (ENSR00000344551) did not reveal *KLF3*- or *KLF15*-binding sites(Fishilevich et al., 2017).

In vitro interventions to modulate adipolin levels have primarily targeted inflammatory responses, on the premise that they are important in the regulation of other adipokines in systemic diseases e.g. metabolic diseases(Ouchi et al., 2011) and osteoarthritis(Conde et al., 2012). For example, in 3T3-L1 adipocytes, adipolin transcript levels were reduced by treatment with palmitic acid (a pro-inflammatory saturated fatty acid elevated in obese individuals), TNFα (a pro-inflammatory cytokine)

and tunicamycin (an antibiotic which induces endoplasmic reticulum stress and pro-inflammatory signalling pathways)(Enomoto et al., 2011). In primary chondrocytes, adipolin transcript levels were reduced by treatment with IL-1β and induced by dexamethasone (an anti-inflammatory glucocorticoid)(Conde et al., 2012).

1.3.4 Adipolin signalling

Studies examining adipolin signalling have been performed either by exposing cultured cells to adipolin or by interrogating differences in pathway activation after experimental manipulation in adipolin-knockout mouse models. A main challenge is that an adipolin receptor has not yet been characterised.

Both adipolin isoforms appear to activate different signal transduction pathways in hepatocytes and adipocytes, with the full-length isoform preferentially activating AKT signalling and the globular isoform preferentially activating p44/42 and p38 MAPK signalling(Wei et al., 2012a). Accordingly, the full-length isoform, via AKT, but not the globular isoform, enhanced insulin-stimulated glucose uptake in adipocytes.

However, a different group showed that treating murine cardiomyocytes with full-length recombinant adipolin activated AKT signalling and that this pathway mediated an anti-inflammatory and anti-apoptotic effect(Takikawa et al., 2020). The authors also induced myocardial infarction on wild-type (WT) and adipolin-knockout mice, finding a suppression of damaged-induced AKT signalling in the hearts of adipolin-knockout mice. In a separate study by the same group, adipolin was found to promote TGFβ receptor II-dependent activation of Smad2 and this resulted in an anti-inflammatory response on mouse macrophages, an antiproliferative effect on human vascular smooth muscle cells and increased TGFβ1 secretion by both cell types(Ogawa et al., 2019). Here, adipolin-knockout mice showed decreased phosphorylation of Smad2 and increased phosphorylation of NF-κB and c-Jun N-terminal kinases in injured vessels when compared with WT mice.

1.3.5 Metabolic effects of adipolin

The positive metabolic effects of adipolin are primarily on glucose clearance. In a study of patients with cardiovascular disease and healthy controls, fasting blood glucose negatively correlated with adipolin levels(Babapour et al., 2020). A similar finding was reported on the association between adipolin and blood glucose in a multiple regression analysis of patients with PCOS and controls(Tan et al., 2014a). Increasing blood adipolin in mice by adenoviral expression and recombinant protein administration led to lower blood glucose and improved insulin sensitivity in three different mouse models – lean WT, diet-induced obese WT and leptin-deficient *ob/ob* mice(Enomoto et al., 2011; Wei et al., 2012b). These effects correlated with enhanced insulin signalling in adipose tissue and the liver (enhanced tyrosine phosphorylation of insulin receptor substrate-1 and AKT), but not in skeletal muscle. Independently of insulin, adipolin suppressed hepatic gluconeogenesis by decreasing expression of the gluconeogenic genes *G6pc* and *Pck1* in the liver and increasing adipocyte glucose uptake(Wei et al., 2012b).

However, the findings from adipolin knockout studies are more challenging to interpret. Ogawa et al(Ogawa et al., 2019) reported no difference in body weight or fasting blood glucose levels between WT and homozygous adipolin-knockout mice (aged 12 weeks)without distinguishing sexes, whereas Tan et al(Tan et al., 2020) found that male (but not female) homozygous adipolin-knockout had impaired fasting glucose levels.

The effect of adipolin on fat metabolism is less marked and inconsistencies exist between studies. Overexpression of adipolin led to no observable difference in non-esterified fatty acid levels in WT, diet-induced obese WT or leptin-deficient *ob/ob* mice(Enomoto et al., 2011; Wei et al., 2012b). Although Enomoto et al(Enomoto et al., 2011) reported a reduction in adipocyte hypertrophy in adipolin-overexpressing diet-induced obese mice, this was not the case in the study by Wei et al(Wei et al., 2012b). In adipolin-knockout mouse studies, even though detrimental effects on lipid metabolism (reduced lipid tolerance and increased hepatic triglyceride and

cholesterol levels) were observed in male diet-induced obese heterozygous adipolin-knockout mice, these effects were not seen in their homozygous counterparts. Unlike their male counterparts, female heterozygote adipolin-knockout mice subjected to diet-induced obesity showed minimal change in lipid metabolism (increased hepatic very low density lipoprotein and triglyceride secretion but no differences in hepatic triglyceride and lipid levels). Female homozygous adipolin-knockout mice unexpectedly gained less weight compared to their WT littermates when subjected to diet-induced obesity. Moreover, when subjected to a low-fat diet, not only did these mice gain less weight, they exhibited higher insulin sensitivity (Table 1.1)(Tan et al., 2020). Only one human study to date examined the lipid metabolism effects of adipolin – here, epicardial fat thickness and low density lipoprotein-C levels in patients with cardiovascular disease were negatively correlated with adipolin, but no differences in total cholesterol, triglycerides or high density lipoprotein-C were detected(Babapour et al., 2020).

Table 1.1: Summary of metabolic phenotypes of male and female *C1qtnf12* heterozygous (+/-) and homozygous(-/-) mice.

		Male mice		**Female mice**	
		***C1qtnf12*+/-**	***C1qtnf12*-/-**	***C1qtnf12*+/-**	***C1qtnf12*-/-**
Low fat diet	Body weight	–	–	–	↓
	Adiposity	–	–	–	–
	Energy expenditure	–	–	–	–
	Insulin sensitivity	–	–	–	↑
	Lipid tolerance	–	–	–	–
	Liver triglyceride level	–	–	–	–
	Liver cholesterol level	–	–	–	–
High fat diet	Body weight	–	↑	–	↓
	Adiposity	–	↑	–	↓
	Energy expenditure	–	↓	–	–
	Insulin sensitivity	–	↓	↓	–
	Lipid tolerance	↓	–	–	–
	Liver triglyceride level	↑	–	–	–
	Liver cholesterol level	↑	–	–	–

Adapted from Tan SY et al (2020).

1.3.6 Anti-inflammatory effects of adipolin

The adipose tissue of adipolin-overexpressing diet-induced obese mice had less macrophage infiltration and lower transcript levels of the pro-inflammatory cytokines TNFα,

IL1β and MCP1(Enomoto et al., 2011). *In vitro*, treatment of murine peritoneal macrophages and cardiomyocytes with adipolin-conditioned media and recombinant adipolin inhibited the stimulatory effect of lipopolysaccharide and TNFα on the expression of TNFα, IL1β and MCP1(Enomoto et al., 2011; Ogawa et al., 2019; Takikawa et al., 2020).

Homozygous adipolin-knockout mice subjected to experimental vascular injury and myocardial infarction exhibited pathological remodelling and increased local inflammatory responses amongst which included macrophage markers (CD68, F4/80), a total immune cell marker (CD45), TNFα, IL6 and MCP1(Ogawa et al., 2019; Takikawa et al., 2020). Interestingly, local adipolin expression decreased following the vascular injury, potentially suggesting a regulatory involvement in the response to injury. Consistent with this, by overexpressing adipolin in WT mice, both disease phenotypes were attenuated(Ogawa et al., 2019; Takikawa et al., 2020).

1.3.7 Adipolin and fibrosis

There has been no study directly assessing the effects of adipolin on fibrosis to date, although its role in metabolism and/or inflammation may provide a link(Mack, 2018; Ung et al., 2021). Furthermore, in an adipolin-knockout mouse model of myocardial infarction, the knockout mice showed an increased fibrotic infarct area(Takikawa et al., 2020).

Adipolin has been shown to promote TGFβ1 secretion in cultured macrophages(Ogawa et al., 2019). In this study, TGFβ1 acted as an anti-proliferative and anti-inflammatory cytokine acting on vascular smooth muscle cells through TGFβ1/Smad signalling. Extrapolating these findings to fibrogenesis *in vivo* requires further investigation. TGFβ1, also commonly a profibrotic cytokine, may be differentially induced, depending on various factors including concentration, receptor activation(Bobik, 2006; Goumans, 2002; Myoken et al., 1990), temporal dynamics, and the local cytokine environment(Ashcroft, 1999).

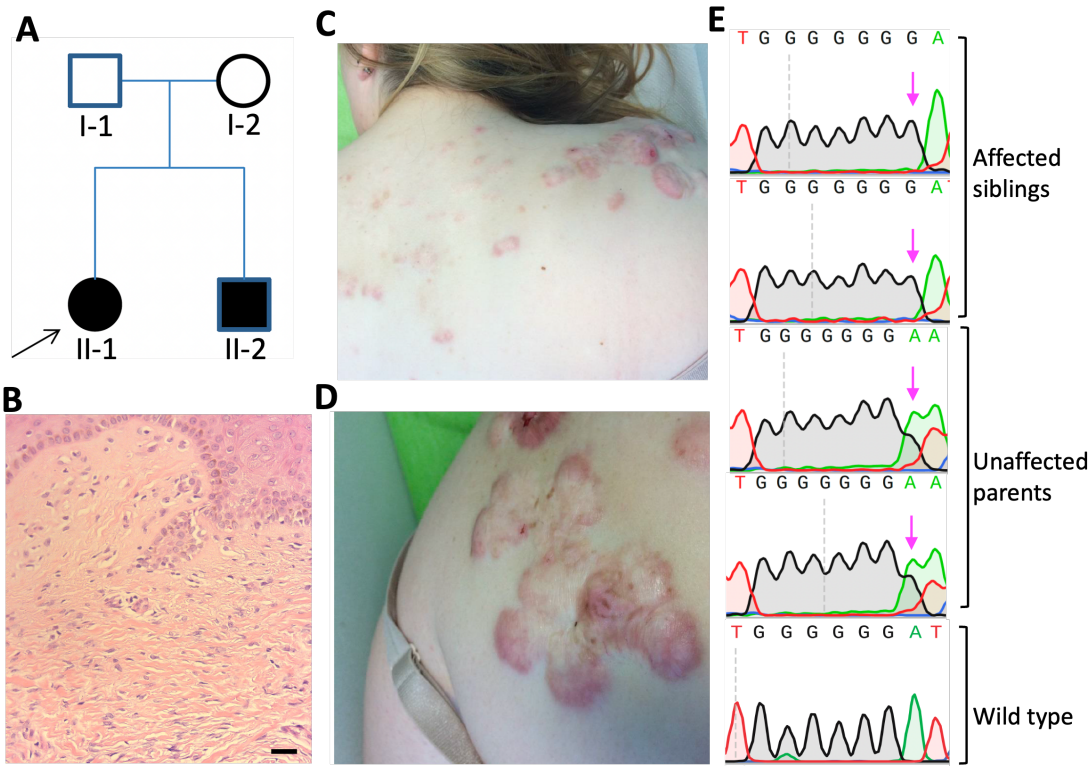

Figure 1.3: Details of excessive scarring pedigree. A. Pedigree chart showing autosomal recessive inheritance pattern. B. Histology of a hypertrophic scar biopsy obtained from the proband. Scale bar = 100μm. C, D. Clinical photographs of proband with severe keloids and hypertrophic scars. E. Sanger sequencing confirmed that the affected siblings in the pedigree are homozygous for the c.126dupG/p.Asn43fs variant (pink arrows) whereas the unaffected parents are heterozygous. The wild-type control sequence is shown.

1.4 Project overview and objectives

This project was conceived following a clinical encounter with a small Israeli family of white ethnicity, whose affected members presented with unusually excessive scarring (both keloids and hypertrophic scars), in an autosomal recessive pattern (Figure 1.3). Through whole exome sequencing, a rare frameshift variant (p.Asn43fs/rs571313759) in the *C1QTNF12* gene (encoding adipolin) that segregated with the phenotype was identified. As the premature termination codon that is introduced is located in the first exon of *C1QTNF12*, this variant likely leads to nonsense mediated decay (mRNA degradation by mRNA surveillance machinery)(Kervestin and Jacobson, 2012), giving rise to a LOF of the gene.

One hypothesis of this finding is that the resultant loss of adipolin function

increases the predisposition to excessive scarring, at a local level by producing a susceptible fibroblast phenotype or at a systemic level, by producing a susceptible metabolic phenotype. The phenotype exhibited by this pedigree is thought to represent an extreme end of the excessive scarring spectrum, serving as a useful model to investigate its genetic and phenotypic basis. An initial focus on adipolin is followed by a phenome- and genome-wide study of individuals with excessive scarring within the UK Biobank (UKB), addressing the following objectives:

- To characterise fibroblasts carrying the *C1QTNF12* LOF variant and to analyse the effects of its protein product, adipolin, on fibroblast proliferation and contractility
- To explore the phenotype of carriers of *C1QTNF12* LOF variants
- To explore the phenotypic and genetic associations of individuals with excessive scarring within the UKB

2 Methods

Contents

2.1 Whole exome sequencing of keloid pedigree

One Israeli pedigree with autosomal recessive keloids was studied with ethics approval and informed consent in collaboration with researchers from Tel-Aviv Sourasky Medical Center, Professor Eli Sprecher and Dr Ofer Sarig. Pedigree identification and sample (blood) collection was performed by researchers from Tel-Aviv Sourasky Medical Center. Whole exome sequencing and Sanger sequencing validation were performed by members of the McGrath lab (King's College London). I performed the whole exome sequencing data analysis.

Whole exome sequencing

Whole exome capture was performed by in-solution hybridization using the Agilent Sure Select Human All Exon 50Mb kit and sequence data was generated on the Illumina HiSeq2500 system (Illumina®).

First, DNA fragments of 150-200bp were obtained by shearing genomic DNA using focused ultrasonication (Adapted Focused Acoustics®, Covaris®). DNA fragments were purified using the Agencourt AMPure XP beads on a Dynal magnetic separator (Invitrogen). The fragment ends were repaired, ligated with sequence adaptors, amplified and hybridised against biotinylated 120bp RNA probes (Agilent), targeting protein-coding genomic regions. Targeted regions were selected for with magnetic streptavidin-coated beads while unbound DNA was washed off. The purified DNA pool was eluted, amplified with low-cycle polymerase chain reaction (PCR), multiplexed (4 samples on each lane) and sequenced with 100bp paired-end reads. Reads were aligned to the Genome Reference Consortium Human Build 37 (GRCh37) using Novoalign (Novocraft Technologies).

Duplicate reads, arising from optical duplicates or PCR clonality, and reads mapping to multiple regions were excluded from downstream analyses. Depth of sequence coverage was calculated using the Bedtools package(Quinlan, 2014) and

custom scripts, whereas variant calling and quality filtering were performed using the SAMtools software package(Li et al., 2009). Variants were annotated with multiple passes through the ANNOVAR software package(Wang et al., 2010).

Polymerase chain reaction

Primers for exon 1 were designed(Ye et al., 2012) to validate the segregation of the *C1QTNF12* variant with the phenotype. PCR was performed according to the AmpliTaq Gold 360 PCR protocol (ThermoFisher Scientific). All samples and reagents were kept on ice. Briefly, a master mix was prepared for each exon containing the following components: AmpliTaq Gold 360 DNA Polymerase (ThermoFisher Scientific) 12.5μl; forward primer (20μM) 1μl; reverse primer (20μM) 1μl and water 9.5μl. 24μl of master mix was added to each well followed by 1μl of DNA (10ng/μl); with a final reaction volume of 25μl. Once DNA was added, the plate was sealed, gently vortexed and centrifuged at 400rpm for 15 seconds. The samples were run on the Thermal Cycler Programme for 1.5 hours under the following cycling conditions: 10 minutes initial denaturation at 95°C, 30 cycles of replication (annealing 15 seconds at 94°C, denaturing 15 seconds at 60°C, and extension 30 seconds at 72°C) and a final extension of seven minutes at 72°C. 3μl of the PCR products were run on a 2% agarose gel (2g agarose, 100ml 1x Tris-borate-EDTA, 2.5μl ethidium bromide) at 100 volts for 15 minutes and the bands of the amplified samples were visualised under an ultraviolet transilluminator.

Sanger sequencing

The PCR products were cleaned with ExoSAP-IT™ PCR Product Clean-up Reagent (Applied Biosystems). 1μl (10ng/μl) of each PCR product was added to 0.25μl of ExoSAP-IT™ and 5.75μl molecular grade water. The mixtures were run on a Thermal Cycler Programme for 30 minutes at 37°C and 15 minutes at 80°C. For the sequencing reactions, the BigDye terminator kit (Life Technologies, Carlsbad, CA, USA) was used to sequence the products. 3.5μl of DNA sample was added to 1.25μl BigDye sequence buffer, 0.25μl BigDye® Terminator v3.1 Ready Reaction

Mix and 0.25μl (20μM) of forward or reverse primer. The forward and reverse primer reactions were run separately. Both reactions were incubated in a Thermal Cycler Programme for 1.5 hours at 96°C for 30 seconds, 50°C for 15 seconds and 60°C for 60 seconds. Products were purified by ethanol precipitation. 20μl of precipitation solution was added to 5.25μl of each sample, incubated for 15 minutes at room temperature and spun at 3000rpm for 30 minutes. The supernatant was discarded and the samples were washed with 100μl per well of 70% ethanol and spun at 3000 rpm for 10 minutes before being dried at 60°C for five minutes. The samples were then resuspended in 10μl Hi-Di™ Formamide (ThermoFisher Scientific) and heated at 95°C for two minutes before being run on the ABI 3730XL Automated Sequencer (Applied Biosystems). Chromatograms were visualised and genetic variants were detected using the Sequencer software (Gene Codes Corporation) and SnapGene Viewer (GSL Biotech LLC).

Whole exome sequencing data analysis

Databases

1000 Genomes Project (1KGP) The 1KGP(Auton et al., 2015) reconstructed genomes of 2504 individuals from 26 populations using low-coverage whole genome sequencing, deep exome sequencing and dense microarray genotyping. A total of more than 88 million genetic variants have been characterised.

Exome Variant Server (EVS) The current EVS data (ESP6500SI-V2) are derived from 6503 samples of European ancestry taken from multiple NHLBI GO Exome Sequencing Project (ESP) cohorts and provides all of the ESP exome variants. The data are publicly available for download for scientific purposes (https://evs.gs.washington.edu/EVS/).

Exome Aggregation Consortium (ExAC) ExAC(Karczewski et al., 2016) is a large-scale exome sequencing project generating sequence data from 60706 unrelated individuals for population genetic and disease-specific studies. The database is commonly used as a reference for filtering and identifying causative variants.

In-house exome bank Genetic data and variation have been collected by the Genomics Group at the Department of Genetics and Molecular Medicine (King's College London) from over 6000 individuals encompassing unselected population controls and individuals with rare genetic diseases, who have consented for their exome data to be used for genetic research purposes.

***In silico* variant analysis and stepwise filtering approach** Variants were annotated according to 1KGP, EVS, ExAC and the in-house exome bank frequencies. Variants were assigned pathogenicity prediction scores including Polymorphism Phenotyping v2 (PolyPhen-2)(Adzhubei et al., 2010), Sorting Intolerant From Tolerant (SIFT)(Ng, 2003), Combined Annotation-Dependent Depletion (CADD)(Rentzsch et al., 2018) and Deleterious Annotation of genetic variants using Neural Networks (DANN)(Quang et al., 2014) scores to assess variants based on their potential deleteriousness and biological functionality.

PolyPhen-2(Adzhubei et al., 2010) is an online prediction tool to assess the functional effects of non-synonymous SNPs based on various features including sequence, phylogenetic and structural information. A score from 0-1 is provided to determine the effect of a substitution. Values lower than 0.5 are predicted to be tolerated, whereas scores closer to one are likely damaging.

SIFT(Ng, 2003) predicts a deleterious amino acid substitution distinguishing between a functionally damaging and a functionally tolerated amino acid. SIFT provides a score ranging from 0 to 1. Values lower than 0.05 mean that changes occur in highly conserved residues and are predicted to be damaging, whereas values over 0.05 are likely tolerated.

CADD(Rentzsch et al., 2018) incorporates multiple diverse annotations using a vector machine-based algorithm into a single statistic termed the C score. The cut-off value used to determine deleteriousness was >15.

DANN(Quang et al., 2014) is a tool that uses deep neural networks to filter likely pathogenicity of alleles. Values above the recommended cut-off of 0.095 or approaching one are considered damaging.

The exome variant profiles of both affected individuals were analysed using the autosomal recessive mode, according to the inheritance pattern of the pedigree. Analyses were implemented in R version 4.1.2.

Filtering was first performed according to allele frequency. As rare variants are more likely to result in rare genetic disorders, variants with a minor allele frequency (MAF) of more than 1% in any population-based database were removed. Variant profiles were then filtered to remove those present in the in-house databases. Variant profiles were further filtered by excluding synonymous variants, non-frameshift deletions and non-frameshift insertions, whilst retaining non-synonymous variants, splice site variants, stop gain variants and frameshift insertions/deletions, as the expectation was to identify a protein-altering variant. Using the autosomal recessive model, variants were filtered to require the presence of at least one homozygous or two heterozygeous protein-altering variants within the same gene that were shared by both affected members.

2.2 Adipolin functional studies

Primary dermal fibroblast culture

Primary homozygous *C1QTNF12* mutant fibroblasts (hereafter referred to as I1253) were a kind gift from Professor Eli Sprecher and Dr Ofer Sarig from Tel-Aviv Sourasky Medical Center, Israel. The primary human WT normal and keloid dermal fibroblasts (NDF, KDF) used were previously harvested by Dr Tanya Shaw (ethics reference 14/NS/1073) and placed in cryogenic storage. The growth medium used was Dulbecco's modified eagle media (DMEM) with high glucose (4.5g/L)

supplemented with 10% foetal bovine serum (FBS), 50μg/ml penicillin, 50mg/ml streptomycin and 2mM L-glutamine. For glucose starvation, the growth medium used contained 1g/L glucose (ThermoFisher Scientific, catalogue number 11054001). Cryopreserved vials of cells were thawed, resuspended in complete culture medium into a T-25 culture flask (Corning) and incubated in a 37°C, 5% CO^2 humidified incubator overnight. The culture medium was replaced the following day and every other day until 80% confluency, at which time the cells were subcultured. Cells for subculture were washed with phosphate buffered saline (PBS) and detached using 0.05% Trypsin in ethylenediaminetetraacetic acid (EDTA) solution, which was neutralised with complete culture medium. Cells were then centrifuged for three minutes at 1200rpm, resuspended and split into two or three flasks as required. For this report, cells used were below passage 10.

Cell-derived matrix (CDM) generation

CDM generation, briefly summarised below, was performed based on a previously published protocol (Kaukonen et al., 2017).

Coverslip preparation Sterile 13mm coverslips were placed in each well of a 24-well tissue culture plate and covered in 70% ethanol for 20 minutes then washed three times in PBS. Coverslips were coated in 0.2% sterile-filtered gelatin for 60 minutes at 37°C then washed three times in PBS. The gelatin was then cross-linked with 1% sterile glutaraldehyde (500μl per well) for 30 minutes at room temperature in a fume hood before washing a further three times in PBS. The cross-linking was quenched with sterile-filtered glycine (1M, 500μl per well) in PBS for 20 minutes at room temperature, followed by three PBS washes. Wells were incubated with complete growth medium for 30 minutes at 37°C and washed a final three times in PBS. Plates were either used immediately or stored for up to four weeks at 4°C.

CDM production 45000 cells/well were seeded and cultured overnight. When cells had formed a confluent monolayer, the culture medium was replaced with fresh medium supplemented with sterile-filtered ascorbic acid (50μg/ml). The ascorbate supplemented medium was replaced with freshly prepared medium every two days for a culture duration of 10 days. Medium was then removed from wells and cells were washed with PBS by gentle pipetting. Coverslips containing cells and CDM were fixed in fresh 4% paraformaldehyde for immunofluorescence staining (described further in Section 2.2.4).

Adipogenic differentiation of cultured dermal fibroblasts

Adipogenic differentiation was performed by Dr Willow Hight-Warburton. Briefly, for a 24-well plate, 40000 cells were plated in 500μl of normal growth medium. At 100% confluence, cells were subjected to three cycles of adipogenic induction and maintenance. Adipogenic induction involves a 3-day culture in dexamethasone (1μM), insulin (10μg/ml), 3-isobutyl-1-methylxanthine (Sigma, 500μM) and indomethacin (Sigma, 50μM) in DMEM supplemented with 10% FBS. This is followed by three days of culture in adipogenic maintenance media consisting of DMEM supplemented with 10% FBS and insulin (10μg/ml). After three cycles of adipogenic induction and maintenance, cells were cultured for a further seven days in adipogenic maintenance media with media replacement every two to three days. Adipogenic differentiation was confirmed by Oil Red O staining.

2.2.1 Functional characterisation of fibroblasts

Assessment of cell morphology

100000 cells per well were seeded in 6-well plates in normal growth medium, then left to adhere and grow in the incubator for up to 48 hours. Cells were fixed in fresh 4% paraformaldehyde at six hours, 24 hours and 48 hours and stained with nuclear and cytoskeletal markers for immunofluorescence imaging as described in Section 2.2.4.

Proliferation assay

Proliferation of I1253 fibroblasts was first assayed on coverslips using Ki67 as a proliferative marker (detailed in Section 2.2.4). The proliferation assay was repeated by counting DAPI-positive cell numbers directly in culture vessels over 48 hours. Subsequent proliferation assays including that of normal fibroblasts treated with adipolin conditioned-media was performed using the DAPI-positive cell count method.

For assaying I1253 fibroblasts, sterile coverslips and 2% gelatin solution were prepared by autoclaving. Coverslips were placed in 24-well tissue culture plates, coated with 2% gelatin by incubating at 37°C for one hour. Wells were washed three times in PBS before subsequent introduction of culture medium with or without supplemental sterile-filtered ascorbic acid (50μg/ml). 5000 cells per well were seeded and left to adhere and grow in the incubator for 48 hours.

The proliferation assay of I1253 fibroblasts was repeated by counting DAPI-positive cells. 100000 cells per well were seeded in 6-well plates in normal growth medium (without supplemental ascorbic acid), then left to adhere and grow in the incubator for 48 hours. Cells were fixed at six hours, 24 hours and 48 hours.

For assaying normal fibroblasts treated with adipolin conditioned-media, 3000 cells per well were seeded in 48-well plates and left to adhere and grow in the incubator for 72 hours. Cells were fixed at six hours and 72 hours.

Stimulation assays

For stimulation of NDF and I1253 fibroblasts with TGFβ1, 45000 cells were seeded per well in 12-well tissue culture plates in 500μl of culture medium with and without the addition of TGFβ1 1ng/ml. Cell lysates were collected after 72 hours. Specifically, wells were washed with ice-cold PBS and the cells were lysed with 100μl of radioimmunoprecipitation assay (RIPA) lysis buffer supplemented with 1:100 protease inhibitor cocktail. Wells were scraped and all liquid containing protein lysates transferred into Eppendorfs before being frozen at -80°C for at least overnight.

For stimulation with commerically obtained adipolin, 45000 cells were seeded per well in 12-well tissue culture plates in 500μl of culture medium +/- TGFβ1 1ng +/- full length recombinant adipolin (Aviscera Bioscience #00392-02-100). After 24 hours in the incubator, cells were visualised using a phase contrast microscope with photomicrographs taken.

For stimulation of NDF with adipolin-conditioned medium (Section 2.2.2) for RNA-sequencing, fresh 24-hour adipolin- and control-conditioned medium were collected, spun down and filtered using a 0.22μm-diameter filter. 200000 cells were seeded per well in 6-well tissue culture plates in 1.5ml of the conditioned medium. RNA purification was performed after 48 hours (Section 2.2.3).

For stimulation of collagen gel-embedded fibroblasts, cells were embedded into collagen gels by thoroughly mixing 500μl of cell suspension at 150000/ml, 200μl of cold collagen I solution (ThermoFisher Scientific, A1048301) and 6μl of 1M NaOH to neutralise and set the gel (to obtain a final cell concentration of 100000/ml), taking care to avoid incorporating bubbles. This suspension was dispensed into a well of a 24-well plate and left for 15-20 minutes at 37°C to set. The collagen gel was then released from the edge of the well with a p20 pipette tip and 500μl of culture media (+/- TGFβ1 1ng/ml) was added. For stimulation of the collagen gel-embedded fibroblasts with recombinant mouse adipolin, the experiment was scaled down to 96-well tissue culture plates due to limited recombinant protein. For all experiments, the plates were imaged at 24-hourly intervals for up to 14 days using the BioRad GelDoc+ imaging system. Gel areas were quantified using ImageJ(Schneider et al., 2012).

2.2.2 Adipolin overexpression and conditioned media preparation

Preparation of adipolin DNA construct

A FLAG-tagged DNA construct containing the *C1QTNF12* cDNA in a pcDNA3.1+/-(K)-DYK vector (clone ID OHu26695D) and a His-tagged subclone in a pcDNA3.1+/-C-6His vector were purchased from Genscript.

Bacterial transformation

TOP10 chemically competent *E. coli* were thawed on ice for 10 minutes before addition of plasmid DNA. 2.5μl (0.5μg) of DNA was added to 25μl of bacterial cells and left on ice to anneal for 10 minutes before being heat-shocked for 45 seconds at 42°C in a water bath. The sample was immediately placed on ice for two minutes. 150μl of Luria-Bertani (LB) broth was added and the sample was incubated at 37°C for one hour in a shaking incubator at 225rpm. The transformation reaction was then plated on an agar plate containing the appropriate resistance antibody for the vector (ampicillin). The plate was incubated at 37°C for 16 hours before collection of bacterial colonies. The single largest colony was picked with a pipette tip and added to 50ml of LB broth containing ampicillin and left in a shaking incubator for 16 hours. Glycerol stocks of the transformed bacteria was prepared by adding 250μl of glycerol to 750μl of the LB broth containing the transformed bacteria. The remaining LB broth was centrifuged at 4000rpm for 15 minutes and the pellet was retained for plasmid DNA purification.

Plasmid purification

Plasmid DNA was purified from the pelleted bacterial cells by alkaline lysis using the Qiagen Midiprep kit according to the manufacturer's protocol. Plasmid DNA was eluted in a high salt buffer (Buffer EB) and quantified using a NanoDrop spectrophotometer (ThermoFisher Scientific).

Generation of adipolin-overexpressing Hek293 cells

Transient transfection of Hek293 cells was performed using polyethylenimine (PEI). One day before transfection, 160000 cells/well were seeded in 12-well tissue culture plates in DMEM supplemented with 10% FBS and 2mM L-glutamine to achieve a confluency of 80% at the time of transfection. On the day of transfection, the culture medium in each well was replaced with 900μl fresh medium and the plate was placed in the incubator whilst DNA-PEI complexes were prepared. For each transfection sample, equal volumes of diluted plasmid DNA and diluted PEI were

prepared separately before mixing and adding to the cells. 1μg of plasmid DNA was used per well and the ratio of DNA to PEI used was 2.5. For transfection of one well, two Eppendorf tubes were prepared: one containing 1μg DNA diluted in 50μl of serum free media (Optimem, ThermoFisher Scientific) and the other with 2.5μg PEI diluted in 50μl Optimem. The tubes were vortexed and spun down briefly, then incubated for 10 minutes at room temperature. The diluted PEI was added dropwise to the diluted DNA while gently mixing and the total transfection mixture (100μl) was added to each well to be transfected. Cells were incubated with the transfection mixture for 12-16 hours before lysing for RNA and protein extraction.

Conditioned media preparation

One day before transfection, 2600000 cells were seeded in 100mm dishes in DMEM supplemented with 10% FBS and 2mM L-glutamine to achieve a confluency of 80% at the time of transfection. On the day of transfection, the culture medium in each well was replaced with 6300μl fresh medium and the plate was placed in the incubator whilst DNA-PEI complexes were prepared. For each transfection sample, equal volumes of diluted plasmid DNA and diluted PEI were prepared separately before mixing and adding to the cells. 15μg of plasmid DNA was used per dish and the ratio of DNA to PEI used was 2.5. For transfection of one dish, two Eppendorf tubes, each containing 15μg DNA diluted in 350μl of Optimem and 37.5μg PEI diluted in 350μl Optimem, were prepared. The tubes were vortexed and spun down briefly, then incubated for 10 minutes at room temperature. The diluted PEI was added to the diluted DNA and the total transfection mixture (700μl) was added to each dish to be transfected. After 16 hours of incubation with the transfection mixture, cells were split at 1:3 ratio and cultured in complete media for 24 hours, after which, conditioned media was collected.

2.2.3 Nucleic acid assays

DNA extraction

Genomic DNA (gDNA) for the I1253 fibroblasts were extracted using the Monarch Genomic DNA Purification Kit (New England BioLabs Inc). Briefly, 1000000 cells were pelleted by centrifugation at 1000 x g for 1 minute and resuspended in 100μl cold PBS. 1μl Proteinase K and 3μl RNase A (solutions from kit) were added and mixed. 100μl Cell Lysis Buffer was added and vortexed thoroughly. The sample was incubated for 5 minutes at 56°C. 400μl gDNA Binding Buffer was added and mixed thoroughly before being transferred to a gDNA Purification Column and centrifuged at maximum speed. The column was then transferred to a new collection tube and 500μl gDNA Wash Buffer added. The tube was centrifuged twice for 1 minute at maximum speed and flow through discarded. gDNA was eluted by adding 35μl gDNA Elution Buffer and centrifuging at maximum speed for 1 minute.

Polymerase chain reaction

Primer pairs were designed for amplifying fragments containing the *C1QTNF12* variant as well as selected published keloid variants using Primer-BLAST (Table 4.3).(Ye et al., 2012) PCR was performed according to the AmpliTaq Gold 360 PCR protocol (ThermoFisher Scientific). All samples and reagents were kept on ice. Briefly, a master mix was prepared for each exon containing the following components: AmpliTaq Gold 360 DNA Polymerase (ThermoFisher Scientific) 12.5μl; forward primer (20μM) 1μl; reverse primer (20μM) 1μl and water 9.5μl. 24μl of master mix was added to each well followed by 1μl of DNA (10ng/μl); with a final reaction volume of 25μl. Once DNA was added, the plate was sealed, gently vortexed and centrifuged at 400rpm for 15 seconds. The samples were run on the Thermal Cycler Programme for 1.5 hours under the following cycling conditions: 10 minutes initial denaturation at 95°C, 30 cycles of replication (annealing 15 seconds at 94°C, denaturing 15 seconds at 60°C, and extension 30 seconds at 72°C) and a final extension seven minutes at 72°C. 3μl of the PCR products were run on a 2% agarose gel (2g agarose, 100ml 1x Tris-borate-EDTA, 2.5μl ethidium bromide)

at 100 volts for 15 minutes and the bands of the amplified samples were visualised under an ultraviolet transilluminator.

Table 2.1: Primers used for DNA amplification.

Variant	PMID	Forward Primer (5'-3')	Reverse Primer (5'-3')
rs571313759 (C1QTNF12)	—	AGGCGAGGGGAGGATTTACT	AGGAGAGGGCTTCCCAGA
rs873549	20711176; 23667473	TCAGAGGACATCACAAGGAGA	GGACAGATGACAGATGGCAAG
rs1511412	20711176	GCCAACCAGACTCACATGTG	ATTGGCAAGGAAGTGGAGCA
rs940187	20711176	CCACTGGGCCTTTGTTTCAAG	CAGCTACCTCCAGCCACTTT
rs8032158	20711176; 23667473	TTAATTTTGTCACGACTGCATGG	CACATACTTTGTTTTCTGCCTCCT
rs192314256	32514122	TCTTGAACTTGACCAGGCCT	CTCAAGGAGGGCGTGCTG

Sanger sequencing

The PCR products were cleaned with ExoSAP-IT™ PCR Product Clean-up Reagent (Applied Biosystems). 1μl of each PCR product was added to 0.25μl of ExoSAP-IT™ and 5.75μl molecular grade water. The mixtures were run on a Thermal Cycler Programme for 30 minutes at 37°C and 15 minutes at 80°C. Samples were submitted to Source BioScience for Sanger sequencing. Chromatograms were visualised using SnapGene Viewer (GSL Biotech LLC).

RNA purification

Cells were lysed in RLT buffer (Qiagen, 350μl per 1000000 cells) and stored at -80°C until purification. Samples were homogenised with the QIAshredder (Qiagen) to maximise yield. RNA was purified using the Qiagen RNeasy mini kit with on-column DNAse treatment to remove genomic DNA contamination. RNA from normal skin and keloid scars used were obtained from Dr Shaw's cryopreserved bank. Homogenised whole 4mm tissue biopsies, including epidermis, tissue-preserved in

RNAlater, were processed using the Qiagen RNeasy Fibrous Tissue Mini Kit. The resulting RNA samples were eluted in RNAse-free water. RNA quantification and purity assessment was performed using a NanoDrop 2000 Spectrophotometer (ThermoFisher Scientific); samples with A260/280 ratios of between 1.9 and 2.1 were deemed acceptable for quantitative PCR (qPCR).

Reverse transcription

Reverse transcription was performed from RNA (described above) to synthesise cDNA using the RevertAid First Strand cDNA Synthesis Kit (ThermoFisher Scientific) using 1μg total RNA, 2μl of 10mM dNTP mix, 1μl of RNase inhibitor (20U/μl), 4μl of 5x reaction buffer and 1μl of reverse transcriptase (200U/μl).

qPCR

Gene expression was assessed using real-time PCR with the Thermofisher Scientific Taqman assay. Hs01012116_g1, which spanned exon 3-4 of the *C1QTNF12* gene, was selected for its best coverage; *GAPDH* (Hs02786624_g1) was used as an internal control. The PCR reaction mix was prepared as per the manufacturer's protocol using 10μl of the gene expression mastermix containing the DNA polymerase, 1μl of the gene expression assay, and 1μl (100ng) of cDNA per well in a 96-well optical reaction plate. Samples were prepared in triplicates. The PCR was run using the Applied Biosystems 7900HT RT PCR system at standard settings for 40 cycles.

Transcriptome sequencing and data analysis

For bulk RNA sequencing, RNA purity and integrity was confirmed using the Agilent 2100 BioAnalyzer (Agilent Technologies), as per manufacturer's protocol. Samples with concentration >30ng/μl and RNA Integrity Number >8 were deemed acceptable for analysis. Total RNA was submitted to BGI Tech Solutions for library preparation and transcriptome sequencing. Briefly, mRNA enrichment and purification were performed by oligo dT selection. Non-stranded libraries were prepared and sequencing was performed on the BGISEQ platform. Paired-end reads of 100 nucleotides were generated.

For transcript assembly, clean RNA reads were aligned to GRCh38 using the HISAT package and to the reference genes using the Bowtie2 package. Transcript quantification was performed using the RSEM package. Differential expression analysis and hierarchical clustering were performed using DESeq2 and the R package pheatmap respectively(Kolde, 2019). P-values were adjusted by the Benjamini-Hochberg method, controlling the False Discovery Rate (FDR) at 0.05. A liberal fold-change >1 was set to detect all significant changes in transcript expression, as weak effect sizes were expected due to the limited sample size and low adipolin dosing in the conditioned media. Data analysis was performed using R software version 4.1.2.

2.2.4 Protein assays

Immunofluorescence and imaging

As described in Section 2.2, cells with and without CDM were fixed with fresh 4% paraformaldehyde in PBS for 10 minutes and permeabilised with 0.1% (without CDM) or 0.2% (with CDM) Triton X-100 in PBS for 10 minutes. The Triton solution was removed and washed twice with PBS before staining was performed. After fixing, coverslips were washed three times with PBS and permeabilised with 0.2% Triton X-100 for 10 minutes before a further three washes with PBS. The coverslips were blocked with 4% bovine serum albumin (BSA) in PBS for one hour. The blocking solution was removed and the matrices were incubated overnight with 100μl of primary antibody (anti-Ki67, 1:500, Abcam) diluted in 4% BSA per coverslip at 4°C.

The primary antibodies were removed and coverslips were washed with PBS before incubation in the appropriate secondary antibodies (e.g. AlexaFluor goat anti-rabbit, 1:1000) and fluorophores (e.g. 4', 6-diamidino-2-phenylindole, DAPI, 1:3000, ThermoFisher Scientific; Alexa Fluor 488^{TM} phalloidin, 1:500, ThermoFisher Scientific) for one hour at room temperature in the dark. The secondary antibodies were removed and the coverslips washed with PBS followed by distilled water before mounting on microscope slides. The slides were left to dry overnight in the dark at room temperature before imaging. Tile scans of each well were obtained with

fluorescent microscopy at 10x and 20x objective using the Evos FL Auto 2 Cell Imaging System.

Standardised thresholding of grey-scale images of DAPI-positive cells was performed before automatic quantification on ImageJ. Fluorescent cells within a standardised area of each well were analysed using a pipeline on CellProfiler(Stirling et al., 2021). Briefly, cell borders were automatically traced to obtain parameters including filtered cell area and cell eccentricity. Cell eccentricity (the ratio of focal length to the major axis length) is a measure of cell shape, whereby a perfect circle has an eccentricity value of 0 while a highly elongated oval has a value approaching 1.

Protein extraction

Total protein was extracted from tissue, cells, and precipitated from conditioned media. Conditioned media was generated and collected as described in Section 2.2.2. One volume of conditioned media (300μl) was added to four volumes of cold acetone (1200μl, -20°C) in Eppendorf tubes, vortexed and incubated for >2 hours at -20°C. The mixture was centrifuged for 10 minutes at 13500 x g. The supernatant was carefully removed and the pellet was dried by evaporation at room temperature for 30 minutes before being resolubilised in 20μl RIPA and 20μl 5x Laemmli sample buffer.

For cell lysate extraction, adherent cultured cells were washed twice with cold PBS before addition of cold RIPA. Cells were then scraped and collected into an Eppendorf tube. The cell suspension was centrifuged at 13500 x g for 15 minutes to pellet cell debris and the supernatant was stored at -80°C for further analysis. Protein extraction from skin was previously generated within Shaw Lab (King's College London) whereas healthy subcutaneous fat protein lysates (from redundant abdominal surgery) were a kind gift by Dr Paul Caton and Dr Elizabeth Evans (School of Cardiovascular Medicine and Sciences, King's College London).

Western blot

4μl of 5x Laemmli sample buffer was added to 16μl of sample lysates (equal volumes assumed equal quantities across conditions). The mixtures were heated for

10 minutes at 92°C. Samples were loaded onto a 15% SDS PAGE (pre-cast Invitrogen NuPAGE) gel and run then transferred to a PVDF membrane for immunoblotting. The membrane was blocked with 5% milk in Tris Buffered Saline (TBS) for one hour at room temperature then incubated overnight at 4°C on a rocking platform with primary antibodies diluted in 0.05% BSA in TBS containing 0.1% Tween and 0.02% sodium azide, as follows: anti-αSMA antibody (1:1000, Abcam, ab21027), anti-GAPDH (1:1000, Sigma, G9545), anti-collagen 1 antibody (1:1000, Abcam, ab292), anti-β Actin antibody (1:1000, Abcam, ab8227), anti-adipolin antibody (1:500, ThermoFisher Scientific, PA5-46452).

After primary antibody removal, the membrane was washed three times in TBS containing 0.1% Tween, then incubated with the appropriate secondary antibodies (goat anti-rat antibody, 1:5000, Abcam, ab7097; goat anti-rabbit antibody, 1:5000, Abcam, ab205718) for one hour at room temperature on a rocking platform. After three washes with TBS containing 0.1% Tween, the blot was chemiluminescently developed with the Pierce ECL HRP substrate to visualise bands. Bands were visualised using ImageLab™ (Bio-Rad).

Immunohistochemistry

Paraffin embedded normal and keloid skin sections used were obtained by Dr Tanya Shaw for studying the molecular mechanisms of tissue repair and scar formation (ethics reference 14/NS/1073). Heat-mediated antigen retrieval with sodium citrate buffer was performed for 20 minutes. The slides were washed twice for five minutes in TBS and 0.025% Triton X-100 on a rocker then blocked in TBS with 10% normal serum and 1% BSA in TBS for two hours at room temperature. The slides were drained and incubated with the primary antibody against adipolin (1:100, ab121791, Abcam) overnight at 4°C in a humidified chamber. After incubation, the slides were rinsed twice for five minutes in TBS and 0.025% Triton. Slides were incubated in 0.3% H_2O_2 in TBS for 15 minutes followed by the HRP-conjugated secondary antibody diluted in TBS with 1% BSA for one hour at room temperature. After rinsing in tap water and haematoxylin counterstaining, slides

were dehydrated and mounted. Slides were visualised using the NanoZoomer S60 Digital Slide Scanner.

Enzyme-linked immunosorbent assay

Enzyme-linked immunosorbent assay (ELISA) for adipolin was performed according to the manufacturer's protocol (Cusabio, CSB-EL008058HU), using kit reagents. Culture supernatants (conditioned media) were centrifuged for 15 minutes at 1000 x g at 4°C and assayed immediately. Sequential dilutions (1:2) of stock standard were prepared. 100μl of standard and sample per well were pipetted into the assay plate, covered and incubated for 2 hours at 37°C. 100μl of Biotin-antibody was added to each well and incubated for 1 hour at 37°C. Wells were washed three times and 100μl of HRP-avidin was added to each well, for a further 1 hour incubation at 37°C. Wells were washed five times before adding 90μl of 3,3 ,5,5 -Tetramethylbenzidine substrate at 37°C, keeping the plate protected from light. Reactions were halted after 20 minutes and the optical density (OD) of each well were detected with a microplate reader set to a wavelength of 450nm. Experiments were performed with 2 technical duplicates and the average value was reported. OD values from the standard dilutions were used to plot a standard curve from which the adipolin concentration in samples were derived using the formula: concentration = (OD - intercept)/slope.

2.2.5 Statistical analyses

Data from cell-based experiments were used to generate means and standard deviations or medians, 25th and 75th percentile values. All cell-based experiments were performed in duplicates or triplicates and assessed using the Student t-test and ANOVA. P-values <0.05 were regarded as statistically significant. The GraphPad software was used to implement most of the statistical tests in cell-based experiments.

2.3 UK Biobank

2.3.1 Study population

The UKB is a population-based cohort study that recruited >500,000 participants (40 - 69 years at recruitment) who attended one of 22 assessment centers across the UK from 2006 - 2010(Sudlow et al., 2015). Rich health-related information is available, including self-reported health conditions, lifestyle indicators, anthropometric and biological measurements. Prospective data linkage to various national health records has also been performed; Hospital Episode Statistics (covering inpatient visits only) are available for the whole UKB cohort and primary care records have recently been made available for nearly half the cohort. Analyses were restricted to individuals with linked-primary care records to limit the possibility of misclassifying participants as unaffected controls due to missing data.

2.3.2 UKB ethics

All participants provided electronically written informed consent. UKB has approval from the North West Multi-Centre Research Ethics Committee (ethics reference 11/NW/0382). This research forms part of approved project number 15147.

2.4 Exploring *C1QTNF12* variants within the UKB

2.4.1 UKB exome sequencing

Exome sequencing of the UKB participants was performed on the Illumina NovaSeq 6000 platform using the IDT xGen Exome Research Panel v1.0 and supplemental probes (targeting ~19396 genes(Van Hout et al., 2020). Exome sequencing and initial quality control were performed by the UKB central team prior to data release. Reads were mapped from GRCh37 to GRCh38 using the functional equivalence protocol as previously described(Szustakowski et al., 2021) and per-sample files were aggregated and joint-genotyped into a single multi-sample project-level variant call format (pVCF) file for all UKB 200K samples. As detailed on the UKB

Data Showcase (https://biobank.ndph.ox.ac.uk/ukb/label.cgi?id=170), no variant- or sample-level filters were pre-applied to the pVCF files - the pVCF files contain allele-read depths (DP) and genotype qualities (GQ) for all genotypes from which variant- and sample-level QC can be calculated and to which analysis-specific filters can be applied.

2.4.2 *C1QTNF12* variant analysis

pVCF exome data for *C1QTNF12* from ~200000 UKB participants (data release October 2020) was extracted using the Swiss Army Knife tool within the UKB Research Access Platform. Variant annotation was performed using the Ensembl Variant Effect Predictor(McLaren et al., 2016). Predicted LOF (pLOF) variants were defined as frameshift, stop-gained, splice-acceptor and splice-donor variants. Primary and secondary care clinical records for individuals homozygous and heterozygous for these variants were extracted using the R package ukbwranglr(Warwick, 2022a) and mapped to phecodes via ICD10 using codemapper(Warwick, 2022b) and Phecode Map 1.2(Wei et al., 2017) (further detailed in Section 2.5.1). QC filtering was applied to remove samples with DP <10 and GQ <20 (i.e. corresponding to a genotype error rate <1%). A phenome-wide association study (PheWAS) was then performed using the R package PheWAS(Carroll et al., 2014), with multivariable logistic regressions to estimate the effect of heterozygosity for *C1QTNF12* pLOF on the risk of each phecode diagnosis, adjusting for age, sex and ethnicity. A Bonferroni-adjusted p-value threshold of $0.05/1518=3.3\times10^{-05}$ based on the 1518 phecodes tested was used to determine statistical significance. Phecodes with <200 cases or controls were excluded. Statistical analyses and data visualisation were performed with R version 4.1.2 and additional packages including tidyverse(Wickham et al., 2019) and targets(Landau, 2021).

2.5 UKB phenotypic association study

2.5.1 Ascertainment of disease status

Excessive scarring status A broad definition of "excessive scarring" was used, which included a diagnosis of keloid or hypertrophic scar, as both conditions are not well-distinguished in electronic health record coding systems. Other large-scale studies (Biobank Japan(Nagai et al., 2017) and Finngen(Kurki et al., 2022)) have also used this broad definition to define keloid code selection. Both types of excessive scarring can be difficult to differentiate(Gauglitz et al., 2010; Lee et al., 2004) and this approach would enable capture of potential keloid misdiagnoses.

Comorbidities/outcomes A systematic search for excessive scarring disease associations was performed on Medline using the following search query: (case-control studies/ or cohort studies/ or (Risk Factors or (comorbidity or Comorbidity) or comorbidities or (Prevalence or prevalence) or association or predispose or risk).mp.) and (hypertrophic scar.mp. or Cicatrix, Hypertrophic/ or ((Keloid or keloid or keloids) not Acne Keloid).m_titl.).

Clinical code selection Disease status was ascertained through the following data sources: self-reported (verbal interview), linked hospital episode statistics [International Statistical Classification of Diseases and Related Health Problems, ninth and tenth revisions (ICD9 and ICD10)], cancer register (ICD9 and ICD10) and primary care records (Read2 and Read3). Clinical codelists were manually curated for excessive scarring, vitamin D deficiency and atopic eczema (Table 2.2, Table A.1). As an example, for excessive scarring, cases were identified through linked hospital episode statistics and primary care records only; scarring was not a diagnosis included for self-reporting (`https://biobank.ndph.ox.ac.uk/ukb/coding.cgi?id=6`). Only codes obtained through primary care records (Read2 and Read3) were specific to either hypertrophic scar or keloid, whereas those obtained through

linked hospital episode statistics (ICD9 and ICD10) encompassed both hypertrophic scar and keloid (Table 2.2).

For conditions that have pre-existing manually-curated codelists (hypertension and uterine leiomyoma), Read2 and ICD10 clinical codelists were obtained from the CALIBER portal(Kuan et al., 2019), an open-access resource that provides manually-curated clinical codelists for >300 health conditions. These clinical codelists (Table A.1) were minimally adapted and mapped to Read 3 and ICD9 equivalents respectively, using the mapping files provided by UKB Resource 592 (https://biobank.ndph.ox.ac.uk/ukb/refer.cgi?id=592).

Table 2.2: Clinical codelists for keloid and hypertrophic scars.

Disease	ICD9	ICD10	Read2	Read3
Hypertrophic scar	7014	L910	M2y11	X78TS
Keloid	7014	L910	7L19A, M214., M218.	M214., XaC07, XaPxn

PheWAS The PheWAS catalogue with phecodes that represent a single phenotype and groupings that categorize these codes into diseases, signs and symptoms, injuries, poisonings, procedures and screening codes was used(Denny et al., 2013; Wei et al., 2017). Linked hospital episode statistics and cancer register ICD10 records were mapped to phecodes using Phecode Map 1.2(Wei et al., 2017). Linked hospital episode statistics and cancer register ICD9 records were mapped to ICD10 codes using mapping files provided by UKB Resource 592 (https://biobank.ndph.ox.ac.uk/ukb/refer.cgi?id=592) and subsequently mapped to phecodes. Primary care records were similarly mapped from Read2 and Read3 to ICD10 codes, then to phecodes. Read2-to-ICD10 mappings were restricted to those with icd10_code_def '1', '3', '5', '7', '8' or '15'. Read3-to-ICD10 mappings were restricted to those with mapping_status 'E', 'G' or 'D', refine_flag 'C' or 'P', element_num '0' and block_num '0'. Individuals were considered a 'case' if they had at least one occurrence of an ICD10 code mapping to a phecode. An example mapping to the phecode for 'Uterine leiomyoma' is shown in Table A.2.

2.5.2 Statistical analysis

Descriptive statistics are presented as frequencies with percentages for categorical variables and means with standard deviations for continuous variables. For comparisons between categorical variables, Pearson's Chi-squared test was used whereas for continuous variables, Welch's t-test was used. Statistical significance was set at $p<0.05$.

Significant comorbidity associations were further examined using multivariable logistic regression analyses to estimate odds ratios (OR) and the associated 95% confidence intervals (CIs). In each model, excessive scarring was considered the exposure variable, and the comorbidity was considered the outcome. As we were testing four primary comorbidities (hypertension, uterine leiomyoma, vitamin D deficiency, atopic eczema), a Bonferroni-adjusted p-value threshold (0.05/4=0.0125) was used to determine statistical significance when testing whether the excessive scarring effect size was significantly different from zero. Multiple imputation was used to account for missing data (20 imputed datasets generated with the R software package Multivariate Imputation by Chained Equations, mice(Buuren and Groothuis-Oudshoorn, 2011)).

A "minimal model" adjusting for age, sex (except uterine leiomyoma, which was restricted to female participants only, N=125771), and ethnicity and a "full model" adjusting for the additional potential confounding covariates were fitted for each disease association. The additional covariates for each fully-adjusted model were: (1) hypertension: BMI, Townsend Deprivation Index, smoking status, diabetes, hyperlipidaemia(Bozkurt et al., 2016) (2) uterine leiomyoma: BMI, Townsend Deprivation Index, smoking status(Qin et al., 2020; Stewart et al., 2017) (3) vitamin D deficiency: BMI, Townsend Deprivation Index, smoking-status(Holick et al., 2011; Sohl et al., 2014) and (4) atopic eczema: Townsend Deprivation Index, BMI, allergic rhinitis, asthma(Nutten, 2015).

To assess whether the disease associations varied by ethnicity, fully-adjusted multivariable logistic regressions were fitted stratifying by ethnicity for the three

largest ethnic groups. These comprised self-reported white participants (85.9%), black or black British participants (henceforth "black participants", 6.4%) and Asian or Asian British participants (henceforth "Asian participants", 5.2%).

For the PheWAS, multivariable logistic regressions were performed to estimate the effect of excessive scarring status on the risk of each phecode diagnosis, adjusting for age, sex, ethnicity, smoking status, BMI and Townsend Deprivation Index. A Bonferroni-adjusted p-value threshold of $0.05/1518=3.3\times10^{-05}$ based on 1518 phecodes tested was used to determine statistical significance. Phecodes with <200 cases or controls were excluded.

All statistical analyses were performed with R version 4.1.2. Specific R packages included mice(Buuren and Groothuis-Oudshoorn, 2011), gtsummary(Sjoberg et al., 2021), tidyverse(Wickham et al., 2019), flextable(Gohel, 2022), PheWAS(Carroll et al., 2014), targets(Landau, 2021), ukbwranglr(Warwick, 2022a) and codemapper(Warwick, 2022b).

2.6 UKB genetic association study

2.6.1 UKB genotyping and quality control

Genotyping and initial quality control were performed by the UKB central team prior to data release. Genotyping was performed using the Affymetrix UK BiLEVE Axiom Array (807411 markers on 49950 individuals) and the Affymetrix UKB Axiom Array (825925 markers on 438427 individuals)(Bycroft et al., 2018). A variety of QC approaches were used to account for the large cohort size, population structure, and batch-based genotype calling (required for handing the large dataset), as detailed in `https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/genotyping_qc.pdf`. Poor quality markers were identified and excluded by testing for batch effects, plate effects, departure from Hardy-Weinberg equilibrium, sex effects, array effects, and discordance across control replicates. These marker-based QC tests were performed on the largest predicted ancestral subset within the cohort (European, 463844 individuals) to attenuate population structure effects. After filtering sample

outliers, multi-allelic SNPs and SNPs with MAF <1%, genotype imputation (a process of inferring genotypes that are not directly assayed to increase the number of SNPs that can be tested for association) was performed by estimating haplotypes for the full cohort and combining the inferred haplotypes with the Haplotype Reference Consortium, UK10K and 1KGP phase 3 reference panels(Bycroft et al., 2018).

2.6.2 Ascertainment of disease status

Two case definitions were used: excessive scarring (using the same definition in the phenotype study, Section 2.5.1, encompassing keloid and hypertrophic scars) and strict keloid. Whereas ICD codes are non-specific (ICD-10 does not include separate codes for keloid and hypertrophic scar, so that both are indicated by the code 'L91.0'), primary care (Read) codes include Read codes specific to either keloid or hypertrophic scar. Therefore, (strict) keloid cases were defined by individuals with Read codes for keloid ('7L19A','M214.','M218.', 'XaC07' and 'XaPxn') and without Read codes for hypertrophic scar ('M2y11' and 'X78TS'), and (strict) keloid controls were defined by individuals without any of the ICD and Read codes for keloid and hypertrophic scar.

2.6.3 Genetic association

Genome wide association studies (GWAS) typically aim to identify variants (SNPs, single nucleotide changes at specific locations in the human genome sequence, or indels, short insertions or deletions) that are associated with differences in a phenotype within an individual. The presence of related individuals in study populations can lead to biased estimates of variant effect sizes due to non-random sharing of variants between relatives. Early methods account for confounding by relatedness by excluding one member of each pair of related individuals based on SNP-derived relatedness(Purcell et al., 2007; Yang et al., 2011). However, this can lead to reduced study power. Mixed-effects models circumvent this by testing for associations between each genetic variant and the phenotype, conditioning on the sample structure inferred from all genome-wide SNPs(Aulchenko et al., 2007; Jiang et al., 2019; Kang

et al., 2010; Segura et al., 2012; Yu et al., 2005). REGENIE is a computationally-efficient method that implements a logistic mixed-effects model, to model genetic relationships between individuals as a random effect, using the genome-wide Firth logistic regression test(Mbatchou et al., 2021). The Firth correction is a bias-reduction method that allows for the analysis of binary (case-control) traits with unbalanced case-control ratios, by adding a penalty term to the logistic regression test when case-control imbalance exists (e.g. small numbers of cases and large numbers of controls, as we see for excessive scarring)(Firth, 1993).

GWAS was conducted in the UKB using REGENIE (v2) to assess the genetic associations for both ambiguous and strict keloid definitions. Sample QC excluded individuals with excess population-adjusted genotype heterozygosity (the fraction of non-missing markers that are called heterozygous, which may be an indicator of DNA contamination), excess missing call rates (>5%), sex-discordance (mismatch between reported sex and inferred sex based on genotype, which may indicate sample mishandling) and excess 'relatedness', defined as having 10 or more third degree relatives. To reduce confounding due to differences in ancestral background, primary analyses were restricted to individuals with white British ancestry, defined by a combination of self-reported ethnic background and genetic information(Bycroft et al., 2018). Analyses were further restricted to individuals who had not withdrawn from the study by May 2020 and who had linked-primary care records (to limit the possibility of misclassifying participants as unaffected due to missing data). A subset of black participants, again defined by a combination of self-reported ethnic background and genetic information, was selected for candidate variant association testing.

Phenotypic and genetic data was extracted and cohort selection was performed using Python and Spark on the UKB Research Access Platform. REGENIE was accessed through the Swiss Army Knife tool on the Platform. In Step One of REGENIE (i.e. prediction of individual trait values based on genetic data), a subset of linkage-disequilibrium (LD)-pruned SNPs (r^2<0.9) SNPs with MAF>0.01% and

missing genotype call rates <10% that did not deviate from Hardy-Weinberg equilibrium (HWE p-value $> 1 \times 10^{-15}$) were used to calculate a polygenic prediction value for each individual.

Association testing was performed in Step Two, including the polygenic predicted value from Step One as a covariate in the regression model, in addition to other covariates, namely, sex, genotyping array and the top five ancestry-informative PCs derived from genetic data. QC filtering for both Step One and Step Two of REGENIE were performed using PLINK v2 within the Swiss Army Knife tool.

Post-GWAS data analysis and visualisation were performed using R version 4.1.2 and specific packages including tidyverse(Wickham et al., 2019), qqman(Turner, 2018), biomaRt(Durinck et al., 2009), corrcoverage(Hutchinson and Wallace, 2019), phenoscanner(Kamat et al., 2019), coloc(Giambartolomei and al, 2014), GenomicRanges(Lawrence et al., 2013), seqminer(Zhan and Liu, 2015), and flextable(Gohel, 2022).

2.6.4 Inflation factors

Population and cryptic relatedness can cause spurious associations in GWAS. To estimate the extent of this, genomic inflation factors, lambda GC, were computed (using the median of the resulting chi-squared test statistics divided by the expected median of the chi-squared distribution) with a value of one indicating no systemic bias. This was corroborated by using LD score regression (LDSC) whereby the intercept term from the regression is one plus a term proportional to inflation from population stratification and cryptic relatedness (i.e. one indicates no inflation from population stratification)(Bulik-Sullivan et al., 2015b) and by visualising quantile-quantile (QQ) plots of the distribution of observed versus expected p-values from the analysis.

2.6.5 Heritability

SNP-heritability was calculated using LDSC, based on north-west European LD scores(Bulik-Sullivan et al., 2015a).

2.6.6 Meta-analysis

FinnGen comprises 414262 participants with phenotypes derived from ICD (version 8, 9 and 10) diagnosis codes; codes used to define hypertrophic scar were L91.0 (ICD10), 7014B (ICD9) and 70130 (ICD8). Unlike the community-based recruitment of UKB (which results in a volunteer bias towards participants being healthier than average)(Fry et al., 2017), Finngen participants were recruited from hospitals, and are coupled with legacy case cohorts of several diseases (resulting in higher case numbers). The unique genetic makeup of the Finnish population, a relatively young, isolated population, means that disease predisposing alleles are found at higher frequencies that are permitted by selection in larger and older outbred populations(Kurki et al., 2022). FinnGen samples were genotyped with Illumina and Affymetrix arrays (ThermoFisher Scientific, Santa Clara, CA, USA). Genotype QC parameters that were set included: call-rate $\geq 90\%$; variant quality score recalibration 'PASS'; quality by depth, QD ≥ 2 for single nucleotide variants; QD ≥ 6 for indels; variant Hardy-Weinberg equilibrium p-value $\geq 1\text{x}10^{-9}$; and allele count ≥ 3. Genotype imputation was carried out using a Finnish population-specific reference panel as described in the following protocol: dx.doi.org/10.17504/protocols.io.nmndc5e, and confidently-imputed variants were defined by INFO > 0.6. Analysis was restricted to unrelated individuals with Finnish ancestry after removing population outliers; this resulted in 1126 cases and 291889 controls. Association analysis was performed using the SAIGE mixed-effects model logistic regression method, with sex, age, genotyping batch and 10 PCs as covariates. All genotyping, QC and association testing were performed centrally by FinnGen.

Summary statistics were downloaded from Finngen and the coordinates were lifted over to GRCh37 to match UKB using UCSC Liftover(Hinrichs, 2006). Variants were then matched based on chromosome, position, reference and alternative alleles. Variants that were not consistent between the two studies were excluded.

The meta-analysis was performed in METAL version 2020-05-05(Willer et al., 2010) using a fixed-effects standard error-weighted approach. METAL implements

a weighted Z-score method based on sample size, p-value and direction of effect in each study, and an effect-size based method weighted by the study-specific standard error. The fixed-effects model assumes that the genetic effects are the same across the different studies, providing narrower confidence intervals than the random-effects model. The fixed-effect model was undertaken as only two studies were included for meta-analysis (hence it was not possible to precisely estimate between-studies variance for a random-effects model). An association was considered genome-wide significant at $p < 5x10^{-8}$. For each genome-wide significant association peak, a locus was defined as +/-1Mb from the SNP with the lowest p-value.

2.6.7 Risk loci and lead SNP definition

Genomic risk loci were defined using FUMA v1.3.8 (Functional Mapping and Annotation web-based platform)(Watanabe et al., 2017), an online platform for functional mapping and annotation of genetic variants. Risk loci were defined first by determining lead SNPs. Independent significant SNPs were defined by genome-wide significant p-values ($p < 5x10^{-8}$) and the lack of linkage disequilibrium (LD, square of correlation coefficient between indicator variables of alleles in two loci, $r^2<0.6$). Independent lead SNPs were defined by $r^2<0.1$. To define risk loci, LD blocks of independent significant SNPs that were part of the same lead SNPs were merged followed by merging of physically-overlapping LD blocks or those within a 250kb distance.

2.6.8 Fine-mapping

The identification of causal SNPs with associated loci is challenging. Lead SNPs are those with the most significant p-value and may or may not be causal for the observed associations, whereas the other less significant SNPs may be the true causal variant, may be associated due to LD or may be independently associated. Fine-mapping aims to identify potential causal variants amongst those associated with the trait; for this study Bayesian fine-mapping was performed using the R package 'coloc' version 3.2.1(Giambartolomei and al, 2014).

Bayesian fine-mapping uses posterior probabilities (PPs) of causality to quantify the evidence that a single variant in a locus is causal. An approximate Bayes factor (ABF) is calculated from the effect size and standard error of each variant in its associated locus (lead SNP ± 1Mb), assuming a prior variance on the log odds ratios of 0.04 as originally proposed(Wakefield, 2009). ABFs were rescaled to reflect the posterior probability for each variant being causal. 95% coverage of credible sets (i.e. >95% probability that the causal variant is contained in the credible set) was the accepted minimum threshold.

2.6.9 Functional annotation

Functional mapping and annotation of SNPs from the meta-analysis were performed using FUMA v1.3.8 (Functional Mapping and Annotation web-based platform)(Watanabe et al., 2017) and Phenoscanner v2 in R(Kamat et al., 2019).

Clinical trait mapping — The 95% credible set variants were queried for published GWAS (clinical traits) using Phenoscanner v2. Significant associations were set at $p<1\times10^{-5}$.

Gene mapping — Gene mapping was performed using FUMA positional mapping and expression quantitative trait locus (eQTL, loci that explain variation in gene expression levels) mapping (using FUMA and Phenoscanner). Genomic risk loci in which SNPs were in LD ($r^2>0.6$) with independent significant SNPs (defined in Section 2.6.7) were identified. The maximum distance between LD blocks to merge into a genomic locus was 250kb. The genetic data of European populations in 1KGP phase 3 were used as reference for LD analyses.

FUMA positional gene mapping was performed by mapping all genome-wide significant SNPs at each genomic risk locus to genes within a positional window of <10kb. SNPs were also filtered based on a CADD score >12.37 (the accepted threshold for a deleterious score)(Kircher et al., 2014).

For eQTL mapping within FUMA, SNPs were mapped to genes if there was significant association (Benjamini-Hochberg FDR<0.05) with gene expression in fibrob-

lasts, skin and/or blood in pre-specified datasets (TwinsUK, eQTLGen, MuTHER and GTEx v7).

eQTL mapping was also performed using Phenoscanner(Kamat et al., 2019) whereby the 95% credible set variants were queried for eQTL associations. Significant associations were set at $p<1\times10^{-5}$.

Colocalisation analysis — To further investigate whether there was a single shared variant between keloid/hypertrophic scar-associated variants and those regulating the expression of the mapped genes in skin and blood, Bayesian colocalisation tests were performed using the coloc.abf R package(Giambartolomei and al, 2014). For whole blood *cis*-eQTLs, publicly available summary data from the eQTLGen Consortium portal was downloaded. The eQTLGen Consortium analysis is the largest meta-analysis of blood eQTLs to date and comprises 31,684 samples from a total of 37 datasets. For skin eQTLs, publicly available summary data from GTEx v7 for non-sun-exposed suprapubic skin was downloaded with the exception of eQTLs for the gene ENSG00000232679 (*RP11-400N13.1*), for which summary data from GTEx v8 for non-sun-exposed suprapubic skin was used (as eQTL data for this gene was not available in GTEx v7). For fibroblast eQTLs, publicly available summary data from GTEx v7 for transformed fibroblasts was downloaded. These datasets were filtered for eQTLs associated with the mapped genes in respective tissues at $p<0.05$. For GTEx, *cis*-association data were available for SNPs within 1Mb of the transcription start site(GTEx Consortium, 2017), whereas for eQTLgen, available data was within 1Mb of the centre of the gene(Võsa et al., 2021).

A colocalisation test was performed for each locus where the keloid/hypertrophic scar-associated variants (within 1Mb of the lead SNP, maximising the variants analysed, but noting the caveat that some associations may not have a sufficiently large region on both sides of the SNP to determine colocalisation) were also influencing gene expression (in the respective datasets), using coloc.abf with default parameters and priors. Reported sample sizes for respective tissues and datasets were used. As the eQTLgen summary statistics did not include effect estimates and standard errors, these were calculated from the Z-scores supplied(Adams, CD, 2021; Zhu et al., 2016).

This computed posterior probabilities for five hypotheses: PP0, no association with trait one (eQTL signal) or trait two (keloid GWAS signal); PP1, association with trait one only; PP2, association with trait two only; PP3, association with trait one and trait two by two independent signals; and PP4, association with trait one and trait two by a single shared variant. Evidence of a single shared variant is defined by PP4 >0.75 and colocalisation is defined by $PP_3 + PP_4 > 0.99$ and PP4/PP3 > 5(Guo et al., 2015).

2.6.10 Gene-based association analysis

Gene-based association analyses consider the association between a trait and all SNPs within a pre-defined window around genes, rather than the association of individual SNPs. This approach also enables biological pathway-based analysis, which is based on the premise that genetic bases of complex diseases are a result of gene interactions rather than a single gene effect.

MAGMA v1.08(Leeuw et al., 2015) within FUMA was used to perform gene-based association analyses. The main steps in MAGMA involve mapping SNPs onto genes (if within 10kb of a gene) and computation of gene associations with the trait of interest (SNP-wise model which computes a gene-level test statistic based on mean SNP association). P-values are computed using an F-test with the significance threshold set at p<(0.05/number of protein-coding genes mapped). Using the gene p-values, a competitive gene set analysis, which tests whether genes in curated gene sets and gene ontology terms from MsigDB(Liberzon et al., 2015) (with 4728 curated gene sets and 6166 gene ontology terms) are more strongly associated with the trait than genes outside the gene set, was performed. The p-value significance threshold was set at p<(0.05/number of gene sets tested).

3

Investigating a potential role for adipolin in dermal fibrosis

Contents

3.1 Introduction

The effects of adipolin on wound healing and remodeling are not known, although reports on its modulatory effects on vascular remodeling(Ogawa et al., 2019; Takikawa et al., 2020) and inflammation(Enomoto et al., 2011; Ogawa et al., 2019), hint at this possibility. Fibroblasts are the main cells in the dermis, whose characteristics determine whether or not wounding (and its sequelae) results in fibrosis/scarring(Griffin et al., 2020; Jiang and Rinkevich, 2020). These include proliferation, extracellular matrix (ECM) deposition, and ECM remodeling(Liu et al., 2016; Mukhopadhyay et al., 2005).

Experiments in this chapter aimed to investigate whether the LOF in *C1QTNF12,* encoding adipolin, is contributory to the excessive scarring phenotype observed in our pedigree. *In vitro* analyses were performed to investigate the expression and function of adipolin on dermal fibroblasts. Population-level genetic data was analysed to investigate whether LOF variants in *C1QTNF12* are observed in other people suffering from excessive scarring.

3.2 Identifying *C1QTNF12* as a putative pathogenic gene

The exome sequencing coverage and quality statistics for both affected individuals our the autosomal recessive keloid pedigree are detailed in Table 3.1.

Table 3.1: Exome sequencing coverage and mapping statistics for affected siblings (II-1 and II-2).

Characteristic	II-1	II-2
Total reads	**65,244,895**	**83,878,021**
Mapped to target reads	45,867,958 (70.3%)	58,327,035 (69.5%)
Mapped to target reads plus 150bp	49,644,475 (76.1%)	63,276,576 (75.4%)
Mean coverage	118.6	149.8
Accessible target bases	**33,323,618**	**33,323,618**

Characteristic	II-1	II-2
Accessible target bases 1x	33,168,443 (99.5%)	33,213,884 (99.7%)
Accessible target bases 5x	33,053,467 (99.2%)	33,118,816 (99.4%)
Accessible target bases 10x	32,936,411 (98.8%)	33,026,222 (99.1%)
Target bases 20x	32,563,429 (97.7%)	32,785,245 (98.4%)

A summary of the stepwise filtering approach undertaken using the homozygous (Table 3.2) and compound heterozygous (Table 3.3) models are shown below. Filtering for compound heterozygous protein-altering variants returned 13 variants in six genes: *CCDC88B* (rs146686061 and rs201629177); *FAM47A* (rs17855514 and novel c.T1555C variant); *IGFN1* (rs138698246 and rs368038036), *NUDT* (rs375826824, rs201969564 and rs371803324); *SOX30* (rs182220520 and rs138390114); *ZNF790* (rs199677584 and rs201553532)(Table 3.5), mostly missense variants that had been reported before in dbSNP. A review of the biology, tissue expression pattern and disease associations of these genes using the GeneCards database did not reveal obvious links to keloid pathobiology(Stelzer et al., 2016).

Filtering for homozygous protein-altering variants returned two variants (Table 3.4); one of which, rs571313759, in *C1QTNF12*, was frameshift, whereas the other, rs7263910, in *SEMG1*, was missense. As the rs571313759 variant was located within the first (of eight) exons of *C1QTNF12* and introduced a premature termination codon, the likelihood of it leading to nonsense mediated decay (and hence LOF) was high(He and Jacobson, 2015; Lykke-Andersen and Jensen, 2015). C1QTNF12 encodes adipolin, a recently identified adipokine with insulin sensitising properties, reported to be associated with metabolically unfavourable profiles (e.g. glucose intolerance, PCOS, cardiovascular disease, as discussed in Section 1.3.3), whereas *SEMG1* is a seminal protein which aids the formation of the gel matrix encasing ejaculated spermatozoa, associated with male infertility(Stelzer et al., 2016).

Taking predicted variant deleteriousness, rarity and putative gene function into consideration (discussed in Chapter 1), the homozygous frameshift LOF variant in

3. Adipolin and fibroblasts

C1QTNF12, was considered the most likely candidate causal variant and selected for functional study.

Table 3.2: Summary of whole exome filtering process – homozygous filter.

Filter	Sibling II-1	Sibling II-2
Total variants	26,871	26,757
Variants with MAF<0.01 in public databases	2,065	1,997
Variants absent from in-house exome databases	516	478
Homozygous variants	17	15
Homozygous nonsynonymous, splice-site, or insertion/deletion variants	11	10
Homozygous nonsynonymous, splice-site, or insertion/deletion variants shared between both affected siblings	2	2
Nonsense variants	1	1

Table 3.3: Summary of whole exome filtering process – compound heterozygous filter.

Filter	Sibling II-1	Sibling II-2
Total variants	26,871	26,757
Variants with MAF<0.01 in public databases	2,065	1,997
Variants absent from in-house exome databases	516	478
Heterozygous variants	499	463
Heterozygous nonsynonymous, splice-site, or insertion/deletion variants	315	310
Genes with at least 2 mutations	29	25
Shared compound heterozygous variants between siblings	6	6
Nonsense variants	0	0

Table 3.4: Variant profiles resulting from the homozygous filtering process.

Gene	Variant ID	Variant type	CADD	SIFT	PolyPhen2, HumDiv	PolyPhen2, HumVar
SEMG1	rs7263910	nonsynonymous SNV		0.12	0.925	0.932
C1QTNF12	rs571313759	frameshift insertion	22.4	—	—	—

CADD, Combined Annotation-Dependent Depletion score; SIFT, Sorting Intolerant from Tolerant score; PolyPhen2, online prediction tool to assess functional effects of SNPs, trained on all damaging alleles with known effects on the molecular function causing human Mendelian diseases (HumDiv) and all human disease-causing mutations, together with common nonsynonymous SNPs without annotated involvement in disease (HumVar)

Table 3.5: Variant profiles resulting from the compound heterozygous filtering process.

Gene	Variant ID	Variant type	CADD	SIFT	PolyPhen2, HumDiv	PolyPhen2, HumVar
CCDC88B	rs146686061	nonsynonymous SNV	—	0.00	0.185	0.012
CCDC88B	rs201629177	nonsynonymous SNV	25.00	0.02	0.997	0.991
FAM47A	—	nonsynonymous SNV	—	—	—	—
FAM47A	rs17855514	nonsynonymous SNV	—	0.56	0.000	0.000
IGFN1	rs138698246	nonsynonymous SNV	27.50	0.05	0.821	0.740
IGFN1	rs368038036	nonsynonymous SNV	11.46	—	—	—
NUDT19	rs375826824	nonsynonymous SNV	27.40	0.08	1.000	0.981
NUDT19	rs201969564	nonsynonymous SNV	—	0.47	0.846	0.624
NUDT19	rs371803324	nonsynonymous SNV	25.10	0.02	1.000	1.000
SOX30	rs182220520	nonsynonymous SNV	12.08	0.05	0.244	0.086
SOX30	rs138390114	nonsynonymous SNV	24.30	0.05	0.988	0.760
ZNF790	rs199677584	nonsynonymous SNV	—	0.65	0.255	0.060
ZNF790	rs201553532	nonsynonymous SNV	—	0.65	0.001	0.001

CADD, Combined Annotation-Dependent Depletion score; SIFT, Sorting Intolerant from Tolerant score; PolyPhen2, online prediction tool to assess functional effects of SNPs, trained on all damaging alleles with known effects on the molecular function causing human Mendelian diseases (HumDiv) and all human disease-causing mutations, together with common nonsynonymous SNPs without annotated involvement in disease (HumVar)

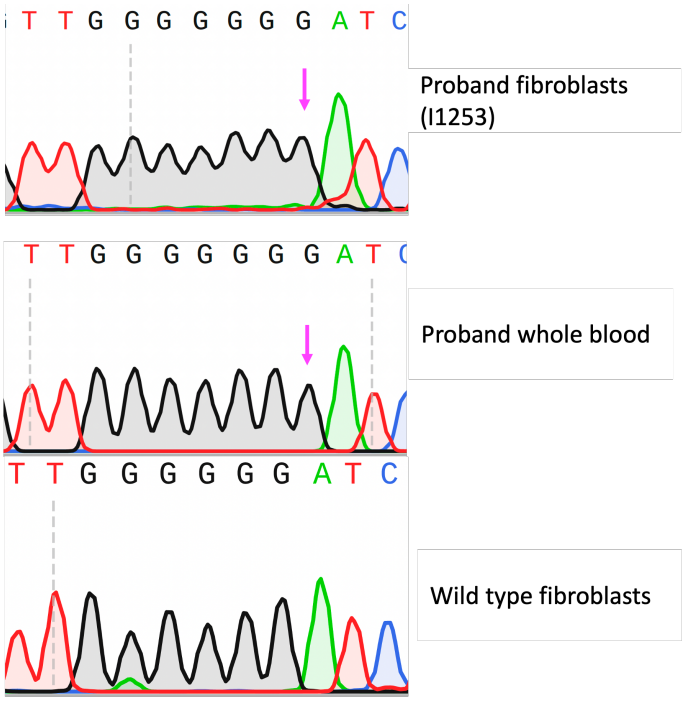

Figure 3.1: Sanger sequencing of cultured fibroblasts from proband confirming the presence of the homozygous c.126dupG/p.Asn43fs variant (pink arrows). The wild-type sequence is shown.

3.3 Characteristics of fibroblasts derived from proband

3.3.1 Genetic sequence

Sanger sequencing was performed on genomic DNA of cultured fibroblasts derived from the proband, I1253, confirming the presence of the p.Asn43fs variant (Figure 3.1). Derivation of primary fibroblasts were performed by the Sarig Lab (Tel Aviv Sourasky Medical Centre) from the patient's upper torso uninvolved skin punch biopsy.

3.3.2 Morphology

Changes in fibroblast morphology may control fibrosis-relevant behaviours such as proliferation(Chen et al., 1997), differentiation(Kilian et al., 2010) and contractility(Alford et al., 2011). The fibroblast shape is a function of its adhesion to its substratum or neighbouring cell(Abercrombie, 1978); elongation correlates with increased traction forces and may promote TGFβ activity(Connor and Gomez, 2013; Wang et al., 2016a). Also, fibroblast size correlates with ECM component production with smaller fibroblasts producing less(Fisher et al., 2015; Quan et al., 2013).

At a qualitative level, there were no observable morphological differences when comparing I1253 fibroblasts to normal WT dermal fibroblasts (NDF), in terms of shape, polarity and size, in standard culture conditions. Quantitative assessment of cellular morphology, conducted by measurement of cell area and eccentricity, confirmed no differences in these parameters between the cell lines compared (Figure 3.2). This was not entirely unexpected as primary fibroblasts isolated from keloid tissue have been observed to exhibit a typical spindle morphology similar in size and shape to normal dermal fibroblasts (NDF), despite being more TGFβ-responsive(Wang et al., 2016b).

3.3.3 Proliferation

It is generally believed that abnormal hyperproliferation of fibroblasts accounts for cell hyperplasia and activation, a feature believed to underlie keloid pathology(Atiyeh, 2020; Zhang et al., 2017). In order to formally determine whether I1253 fibroblasts had altered proliferation, cells were sparsely seeded and stained for the nuclear antigen Ki67 as a marker of proliferative cells(Scholzen and Gerdes, 2000), along with DAPI to mark all cell nuclei. Images and resulting quantitation revealed no significant differences in the proportion of Ki67+ cells. Analyses of cells grown with CDMs (i.e. in the presence of ascorbic acid to stimulate ECM synthesis) over 10 days also showed no observable differences in the proportion of Ki67/DAPI positive

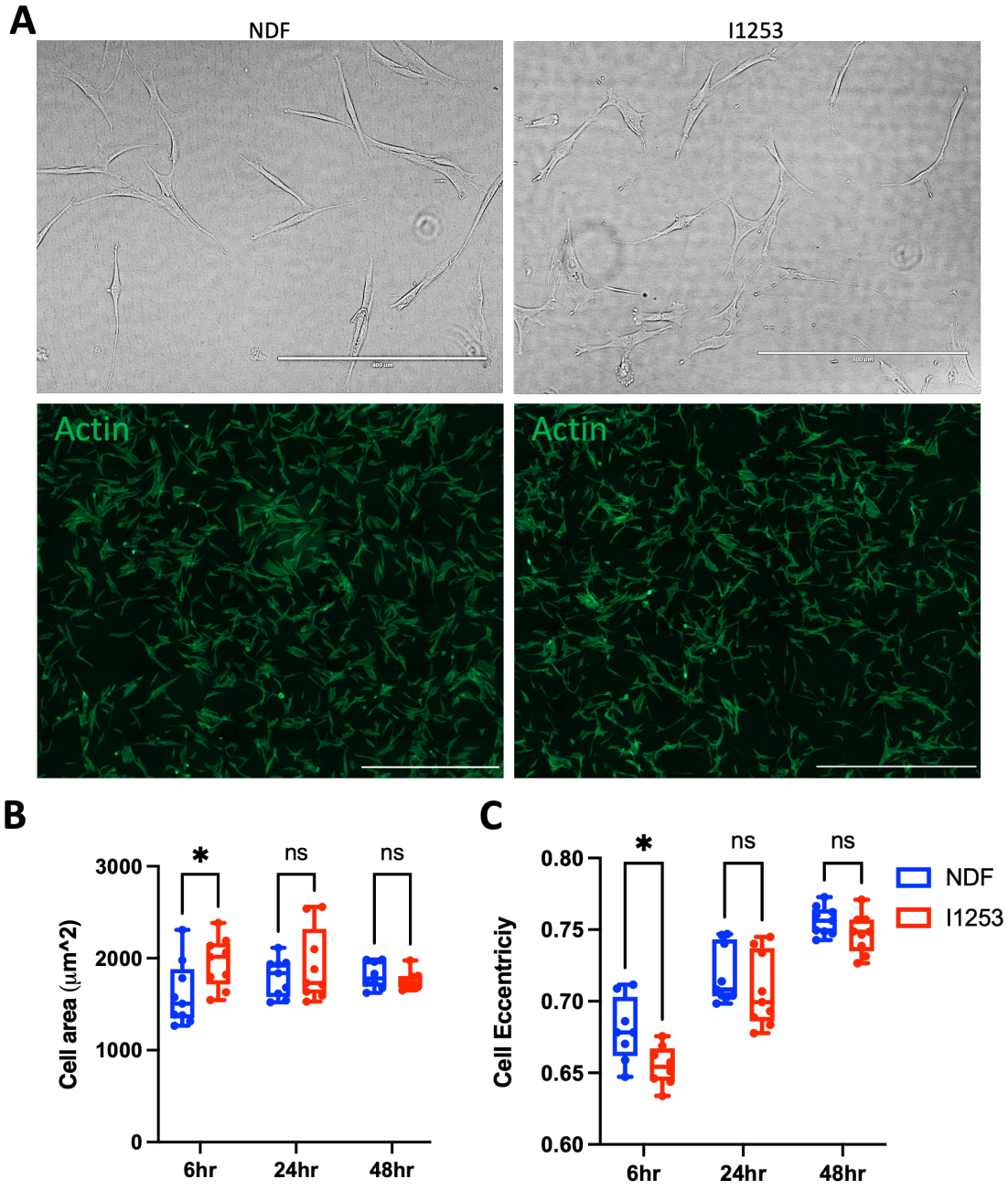

Figure 3.2: Proband (I1253) fibroblasts appear morphologically similar to wild-type normal dermal fibroblasts (NDF). Phase contrast (A, top panels) and actin-stained fluorescent (A, bottom panels) images of NDF (n=2) and I1253 fibroblasts (n=1) under sparse growth conditions on tissue culture plastic. Scale bars: 400μm. B, C. Morphology quantified by analyzing cell area (B) and cell eccentricity (ratio of distance between focus of cell ellipse and its major axis length) with zero representing a circle and one representing a line over 48 hours (C). Each point represents the mean of all cells in one field of view at 20x magnification over two independent experiments. Data are shown as mean ± standard deviation. Unpaired t-tests indicated no significant differences between comparison groups.

cells. An alternative method of measuring cell proliferation, by direct quantification of cell number over time, yielded similar findings (Figure 3.3).

3.3.4 Response to TGFβ1

As a major growth factor in the context of fibrosis, TGFβ1 was selected as the stimulus to elicit profibrotic fibroblast behaviour. KDF are more sensitive to TGFβ1 stimulation, showing increased collagen synthesis(Bettinger et al., 1996) and contractility(Hasegawa et al., 2003). Although derived from unaffected skin, I1253 fibroblasts were hypothesized to be profibrotic, hence may exhibit keloid fibroblast-like properties.

Quite strikingly, although no significant morphological differences were apparent at baseline, the I1253 fibroblasts adopted an enhanced myofibroblast-like appearance following treatment with TGFβ1, with denser aggregations and contractile appearances versus WT fibroblasts, resulting in traction holes within the monolayer (Figure 3.4A)(Evanko et al., 2015).

To investigate whether I1253 exhibited greater contractility in response to TGFβ1, cells were embedded within a free-floating collagen gel supplemented with and without TGFβ1 and monitored over seven days. Consistent with the observed morphological changes, TGFβ1 treatment of I1253 cells led to marked contraction of the gel when compared to NDF treated with TGFβ1 (Figure 3.4B).

Speculating that the increased contractile ability of the I1253 cells is a result of a myofibroblast-like phenotype, western blotting was performed for αSMA, a smooth muscle contractile protein widely used as a marker for myofibroblasts, a major source of pathologic ECM. This showed increased expression of αSMA in I1253 cells and supported the notion that they are profibrotic both at baseline and in the presence of TGFβ1 (Figure 3.4D, Figure A.1).

Hypothesizing that these myofibroblast-like cells may show enhanced collagen synthesis, western blotting was performed, probing for type 1 collagen, a major ECM protein implicated in keloid(Chipev et al., 2000). Interestingly, the I1253 cells

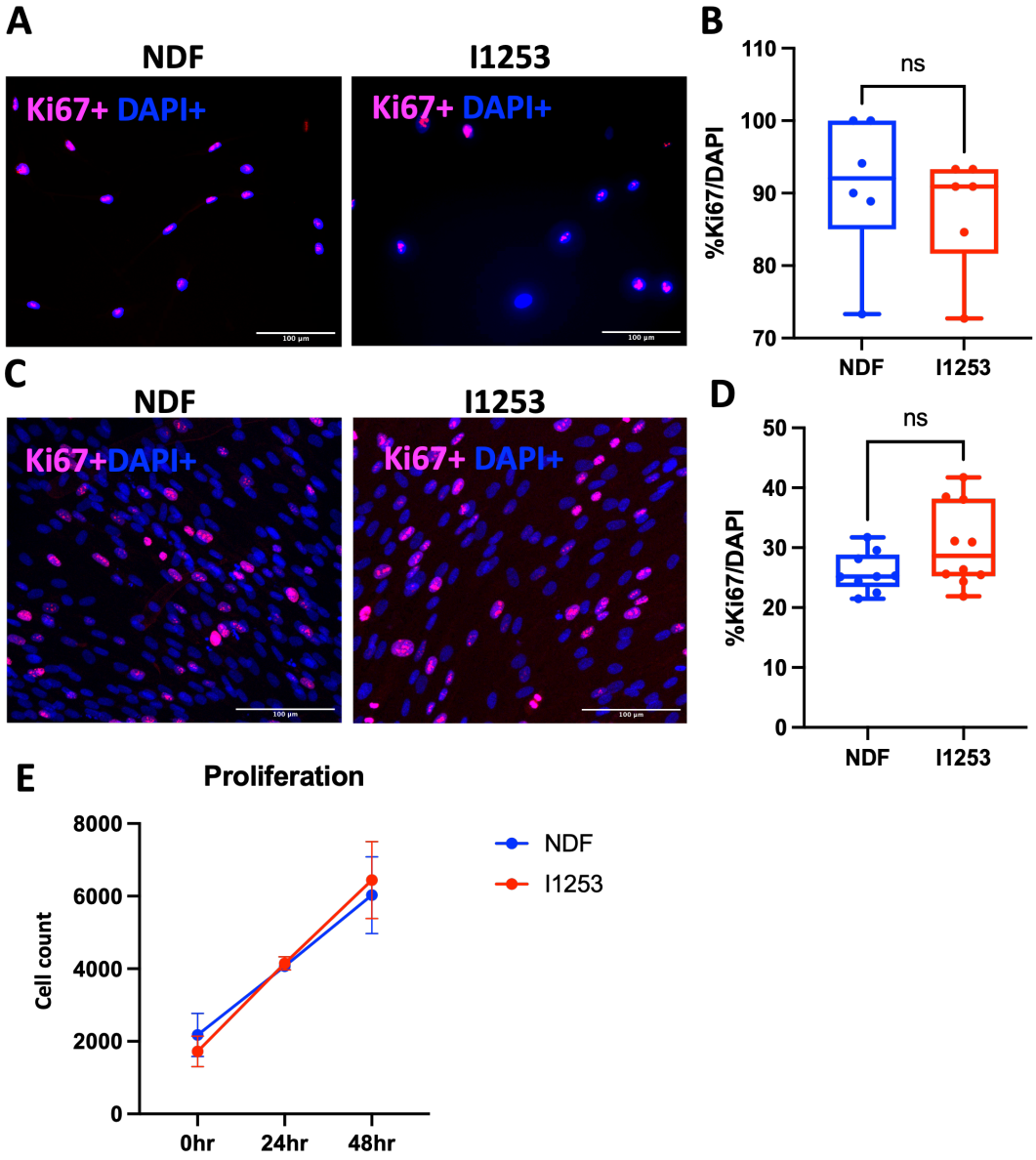

Figure 3.3: Wild-type normal dermal fibroblasts (NDF) and proband (I1253) fibroblasts proliferate at a similar rate. A, C. Example fluorescence microscopy images of NDF (n=2) and I1253 fibroblasts (n=1), fixed and stained for DAPI and Ki67 when cultured for 48 hours in standard 2D culture systems (A) and for 10 days in 3D/cell-derived matrices (C). Scale bars: 100μm. B, D. Proliferation quantified by analyzing the ratio of Ki67 to DAPI positive cells, with each point representing a field of view at 20x magnification. E. Proliferation quantified by direct quantification of DAPI positive cells, fixed and imaged at 4x magnification every 24 hours for 48 hours. Data are shown as mean ± standard error of mean. ns, not significant (unpaired t-test).

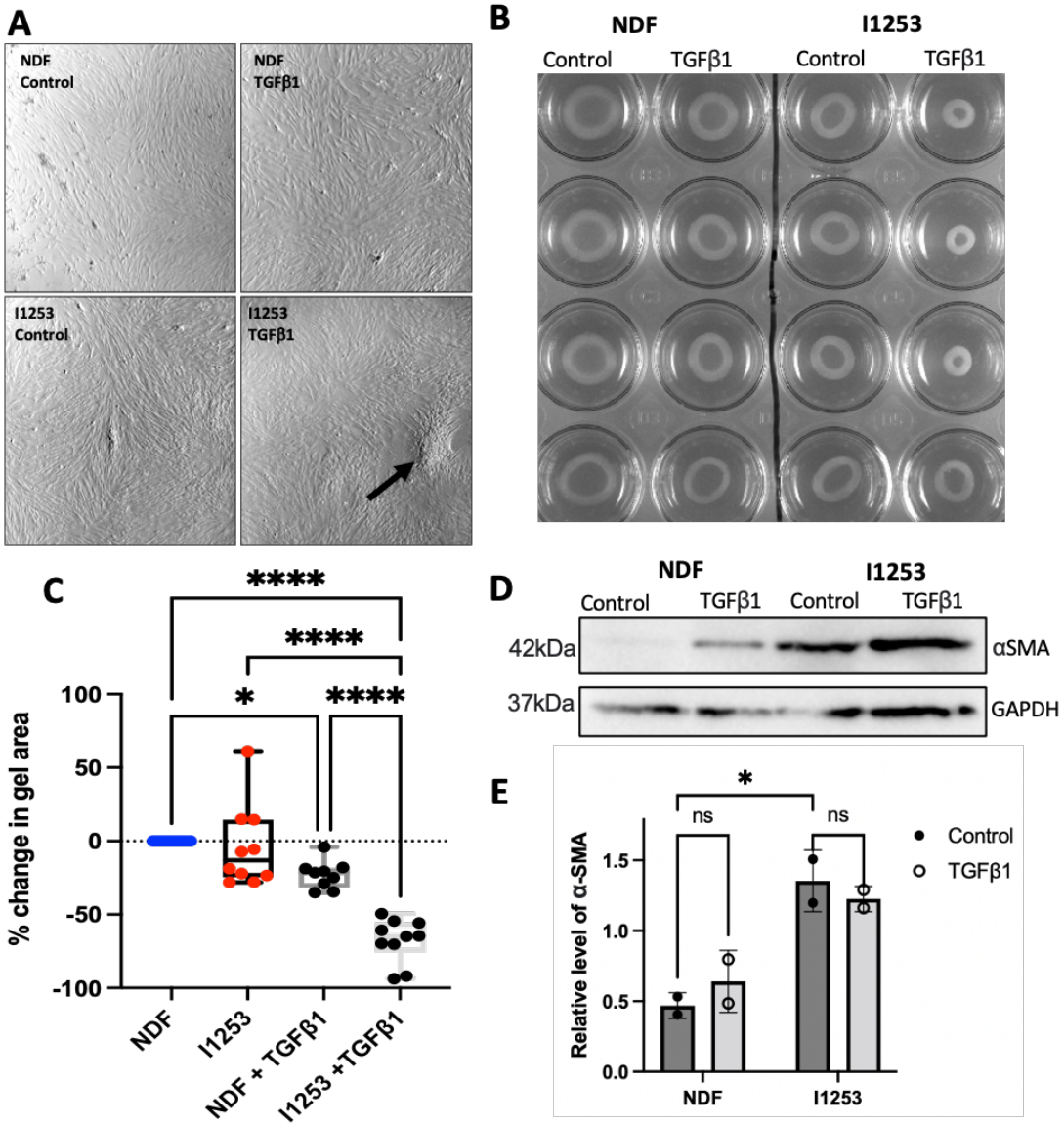

Figure 3.4: Proband (I1253) cells exhibit increased TGFβ1-induced contractility. A. Representative photograph at 4x magnification of normal dermal fibroblasts (NDF) and I1253 fibroblasts at confluency in standard 2D culture conditions. B. Representative photograph of fibroblast-embedded collagen gels at Day Seven. C. Gel area quantified on ImageJ with baseline correction for control gels. Each point represents a gel from a pool of three technical replicates from one biological sample per condition. Pairwise multiple comparisons were performed following 2-way ANOVA with * denoting $P<0.05$ and **** denoting $P<0.0001$. D. Representative western blot of lysates from NDF (n=2) and I1253 fibroblasts (n=1) treated for 72 hours with TGFβ1 1ng/ml. Lysates were probed for αSMA (top panel) and GAPDH as a loading control (bottom panel). E. Densitometry analysis on ImageJ of two representative blots. Pairwise multiple comparisons were performed following 2-way ANOVA with * denoting $P<0.05$.

did not express higher levels of type 1 collagen (Figure 3.5). A representative full western blot image probing for type 1 collagen is shown in Figure A.3.

To summarise, characterisation of I1253 fibroblasts was performed by assessing their morphology, proliferation and response to TGFβ1. Observations of the I1253 fibroblasts indicated no significant differences at baseline apart from αSMA expression, however upon exposure to TGFβ1, they adopted a more "myofibroblastic" contractile phenotype, potentially suggesting a greater propensity to fibrosis.

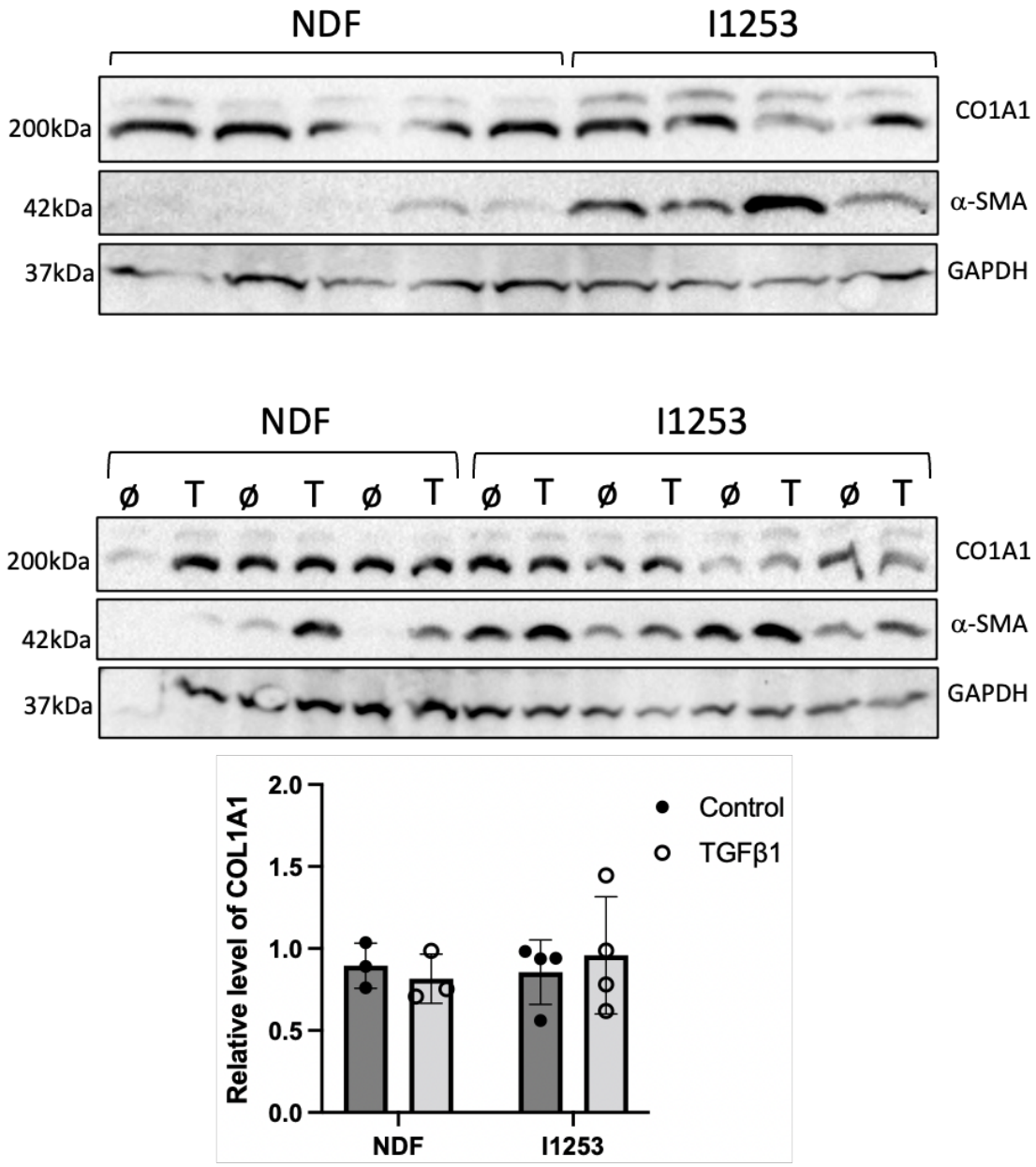

Figure 3.5: Proband (I1253) fibroblasts do not express more type 1 collagen (COL1A1) than normal dermal fibroblasts (NDF). Western blot of lysates from NDF (n=2, with technical repeats at different passages) and I1253 fibroblasts (n=1, with technical repeats at different passages) at baseline (top panel) and ± 72 hour treatment with TGFβ1 1ng/ml. Lysates were probed for COL1A1, αSMA and GAPDH as a loading control (middle panel). A technical replicate for NDF and three technical replicates for I1253 are shown. ϕ, Control; T, TGFβ1. Densitometry analysis on ImageJ (bottom panel). Pairwise multiple comparisons were performed following 2-way ANOVA.

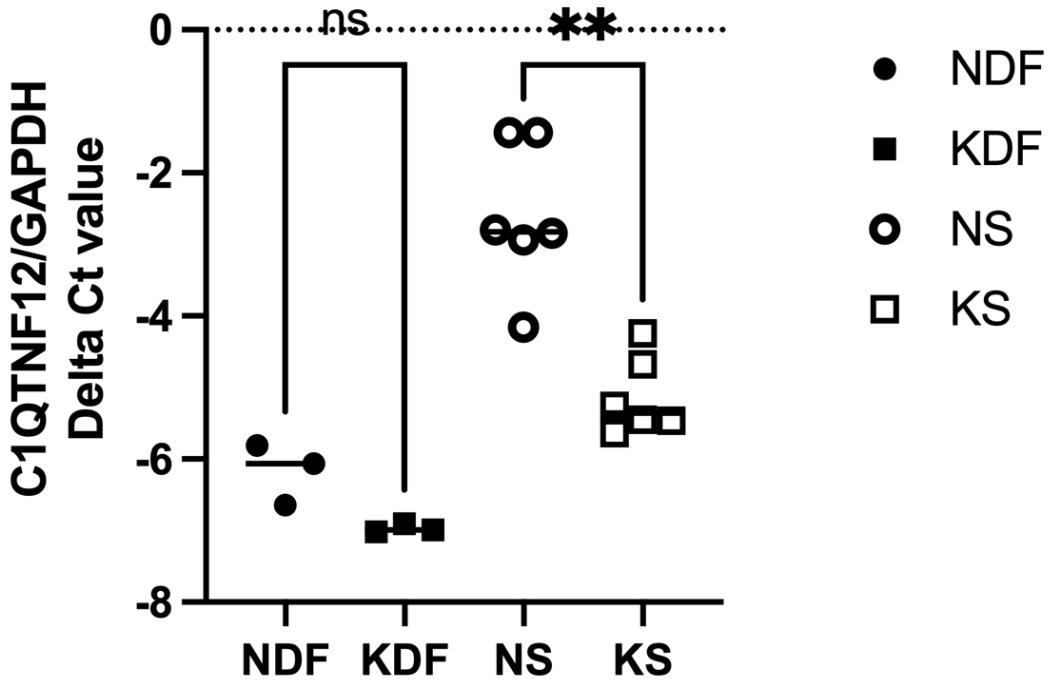

Figure 3.6: Transcript level comparison of *C1QTNF12* for normal dermal fibroblasts (NDF, n=3) versus keloid fibroblasts (KDF, n=3) and normal skin (NS, n=3 with additional technical replicates) versus keloid scars (KS, n=3 with additional technical replicates). Cycle threshold levels in each sample were normalised to GAPDH (delta Ct) and the differences in delta Ct values between groups were analyzed by paired t-test, * indicating $P<0.05$, ** indicating $P<0.01$. Each point represents three technical replicates per biological sample. ns, not significant.

3.4 Expression of adipolin in skin and fibroblasts

3.4.1 Baseline expression

To understand the functional consequences of adipolin in fibroblasts, it was first important to study its normal expression. If adipolin is expressed by WT fibroblasts, its reduced expression would be expected in I1253 (*C1QTNF12*-LOF) fibroblasts. The function of adipolin could then be studied through gene-knockdown/knockout studies, with an aim of rescuing any observed phenotype through its over-expression.

An initial analysis to compare differences in *C1QTNF12* mRNA abundance between NDFs, KDFs, normal skin and keloid scars suggested that it was higher in normal skin compared to keloid scars, but ultimately there was no difference between NDF and KDFs (Figure 3.6).

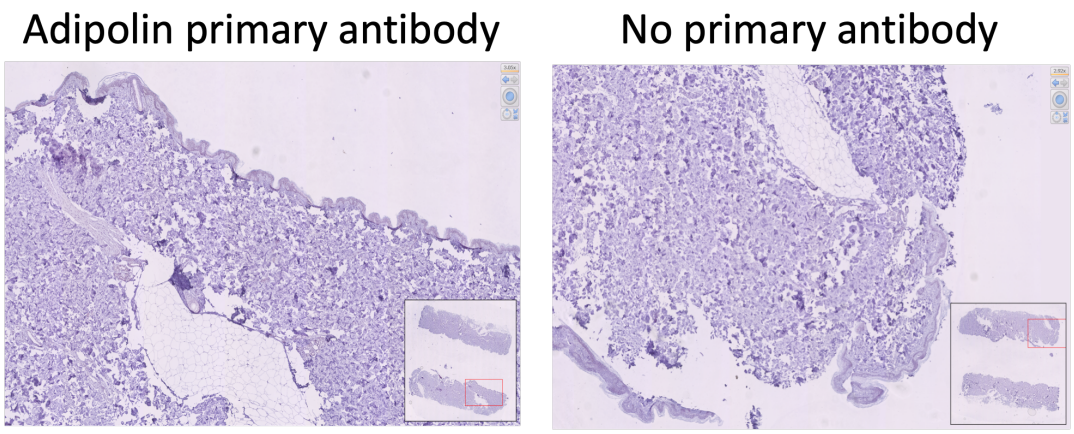

Figure 3.7: Adipolin is not detectable in normal human skin. Representative images shown for normal human skin sections (n=2) stained with primary adipolin antibody plus secondary antibody (left panel) and secondary antibody alone (right panel). A positive signal indicating the presence of adipolin would be brown (following the oxidation of 3,3 - diaminobenzidine, a commonly used chromogen in immunohistochemical staining). The insets show the regions from which the magnified images were obtained.

Hypothesizing that the higher mRNA expression of *C1QTNF12* in normal skin compared to keloid scars was related to the presence of residual subcutaneous fat or pre-adipocytes in normal skin (which would be absent in the fibroblast cultures), immunohisochemical staining of normal human skin sections was performed to question whether the protein was detectable. The skin sections were chosen to include the epidermis and subcutaneous fat, to account for extra-dermal expression, with fat anticipated to be a positive control. As shown in Figure 3.7, adipolin was not detectable using this method/antibody.

Western blotting of a range of cell lines (adipolin-overexpressing HEK293 cells, I1253 fibroblasts, NDF and healthy subcutaneous fat tissue) were probed for adipolin expression. Apart from the adipolin-overexpressing HEK293 cells, no adipolin expression was detected (Figure 3.8A). A representative image of a full western blot image is shown in Figure A.2.

An interrogation of different cell populations in an unpublished integrated single cell RNAseq dataset of healthy skin, non-lesional keloid-prone skin, active and inactive keloid lesions was performed to determine if there were any cutaneous fibroblast populations that expressed the *C1QTNF12* transcript. In particular,

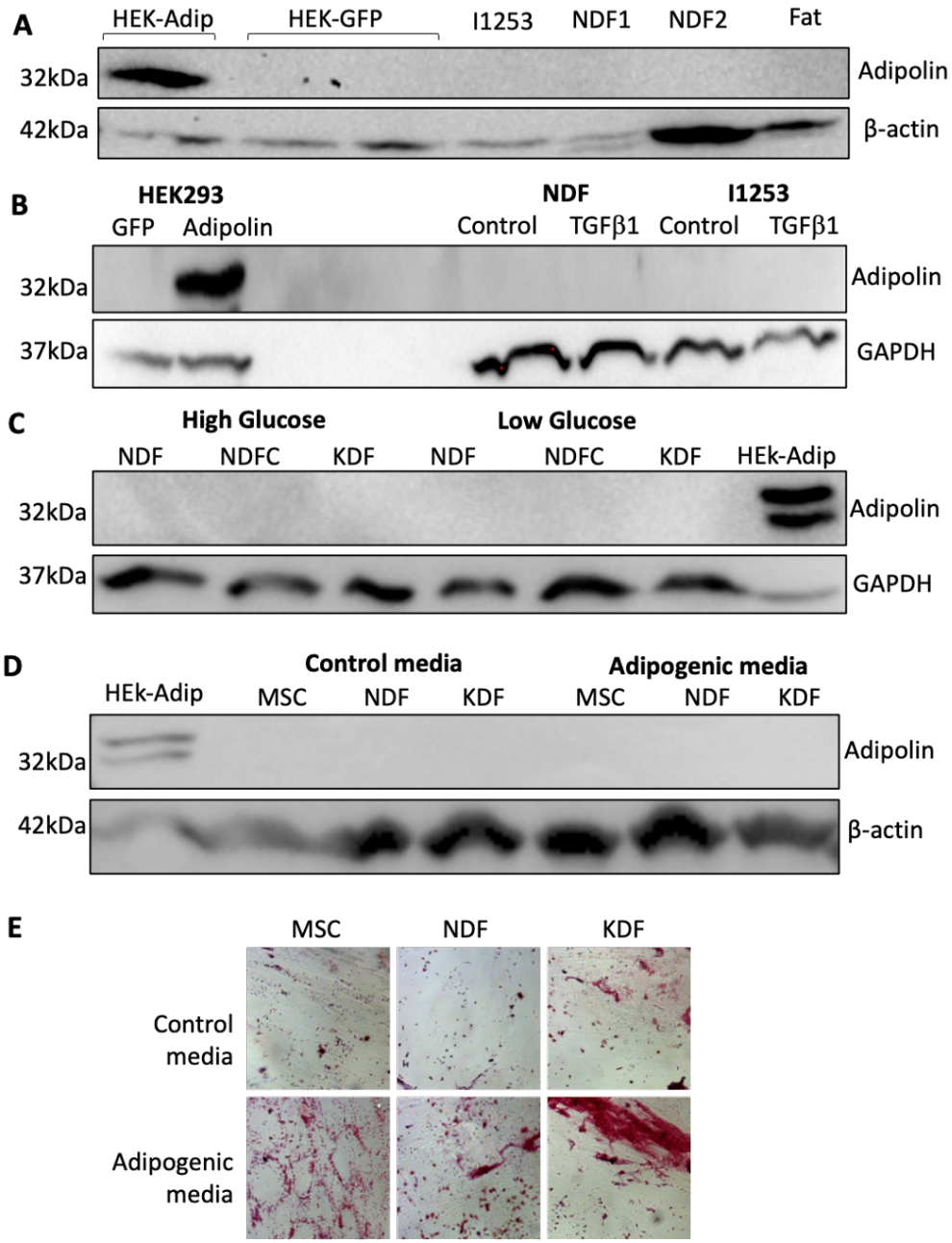

Figure 3.8: Adipolin is not detectable in dermal fibroblasts or subcutaneous fat at baseline or with attempted induction. A. Representative western blots for adipolin of wild-type normal dermal fibroblasts (NDF), proband fibroblasts (I1253) and normal subcutaneous fat at baseline; B. NDF (n=1) and I1253 (n=1) with and without TGFβ1 10ng/ml for 72 hours; C. NDF, keloid-perilesional fibroblasts (NDFC), and keloid fibroblasts (KDF) cultured in standard high (4.5g/L) or low (1g/L) glucose culture media over six days; D. mesenchymal stem cells (MSC), NDF, and KDF with and without an adipogenic differentiation protocol using adipolin overexpressing HEK293 lysates (HEK-Adip) as a positive control. E. A representative image of the adipogenic differentiation of the MSC, NDF and KDF, courtesy of Dr Hight-Warburton is shown.

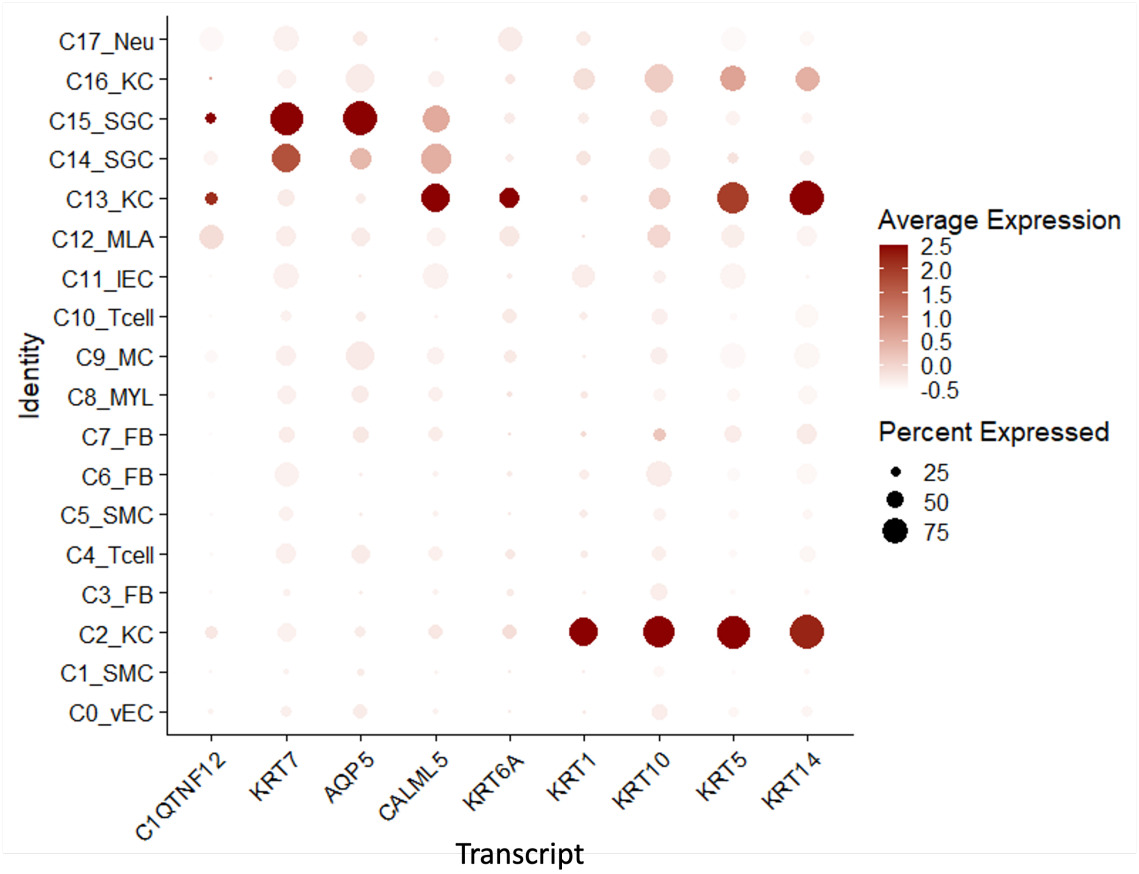

Figure 3.9: *C1QTNF12* transcript was not detectable in fibroblast subpopulations in an integrated single cell transcriptomics dataset of normal, keloid-prone and keloid skin samples (Unpublished data by YK Hong and CK Hsu et al, National Cheng Kung University, Taiwan). FB, fibroblasts; Neu, neutrophils; SGC, sweat gland cells; KC, keratinocytes; MLA, melanocytes; IEC, lymphatic endothelial cells; Tcell, T cells; MC, mast cells; MYL, myeloid cells; FB, fibroblasts; SMC, smooth muscle cells; vEC, vascular endothelial cells.

this was to screen for whether there were pre-adipocyte fibroblast subpopulations that would be enriched for *C1QTNF12*(Shook et al., 2018). This confirmed absent expression in the three major fibroblast subpopulations identified and showed scant expression in subsets of keratinocytes and sweat gland cells (Figure 3.9).

3.4.2 Attempted induction of adipolin expression in fibroblasts

Before concluding absent expression of adipolin in fibroblasts, it was necessary to account for whether this protein was inducible in the context of wounding/scarring

and that the absence of its induction contributed to pathology.

TGFβ1 treatment

To find out whether the differences in I1253 behaviour in response to TGFβ1 were a direct result of adipolin expression (with the hypothesis that WT NDF would upregulate the expression of adipolin in response to TGFβ1 limiting their relative fibrotic response), cells were treated with and without TGFβ1 and probed for adipolin expression. There was no detectable expression as shown in Figure 3.8B.

Glucose starvation

Published data state that at a systemic level, serological adipolin concentration is increased by hypoglycaemic states and decreased by glucose challenges(Tan et al., 2014a). It is also well known that glucose demands are higher in healing wounds(Im and Hoopes, 1970). As standard culture medium for fibroblasts contains high glucose levels (4.5g/L), a variety of primary dermal fibroblast lines (NDF, KDF and keloid-perilesional fibroblasts, NDFC) were grown in low glucose medium (1g/L) and probed for adipolin expression. Western blotting of all cell lysates tested showed no detectable adipolin expression (Figure 3.8C). It is worth noting that two bands are visible for adipolin in the positive control; this may be a result of post-translational modification of the protein, as discussed in Chapter 1.3.1.

Adipogenic media culture

Insulin, an established adipogenic differentiation factor(Scott et al., 2011), is reported to increase adipolin expression in adipose tissue at transcript and protein levels(Wei et al., 2012b). Adipogenesis is postulated to be an anti-scarring process(Hoerst et al., 2019; Plikus et al., 2017) and autologous fat grafting is used to improve scars(Negenborn et al., 2016). Adipolin by definition is expressed by adipose tissue(Enomoto et al., 2011) and although in this study it was not detectable in human subcutanous fat, a consideration was whether reprogramming fibroblasts towards adipogenesis may induce its expression. Cells subjected to adipogenic

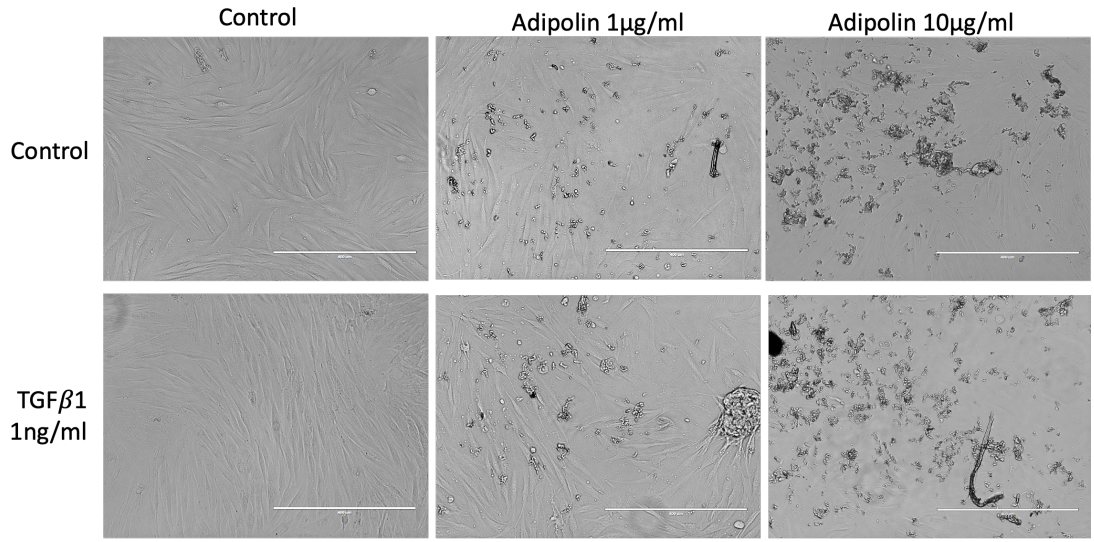

Figure 3.10: Treatment of fibroblasts with commercially-obtained adipolin results in cell death. Phase contrast images of normal dermal fibroblasts untreated (control) or treated with recombinant adipolin (Aviscera Bioscience, product code 00392-02-100, 1μg/ml or 10μg/ml) alone or in combination with TGFβ1(1ng/ml). Scale bars are 400μm.

differentiation (Figure 3.8E) were probed for adipolin but no detectable protein expression was found (Figure 3.8D).

3.5 Effect of exogenous adipolin on dermal fibroblasts

3.5.1 Exogenous adipolin

Human recombinant adipolin purchased from Aviscera Bioscience (product code 00392-02-100, 95% purity, derived from *E. coli*, dissolved in PBS without preservatives) was used to treat dermal fibroblasts at a low dose of 1μg/ml and the published effective concentration of 10μg/ml(Wei et al., 2012b). Within 24 hours, this treatment led to cell death at both concentrations in a dose-dependent manner (Figure 3.10).

As adipolin is reportedly a secreted protein(Wei et al., 2012b), conditioned media from adipolin-overexpressing HEK293 cells was a potential alternative source of exogenous adipolin. To generate adipolin-conditioned media, a *C1QTNF12* overexpression plasmid was purchased from Genscript (CloneID OHu26695, Accession

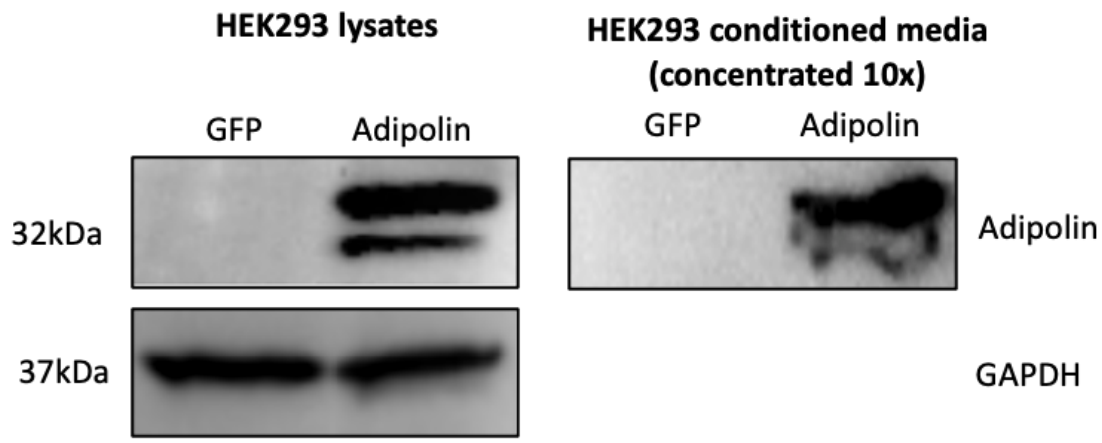

Figure 3.11: Conditioned media of the *C1QTNF12*-overexpressing HEK293 cells contains detectable adipolin. Left panel - western blots of cell lysates from HEK293 overexpressing *C1QTNF12* or GFP. Right panel - Precipitated conditioned media from HEK293 cells overexpressing *C1QTNF12* or GFP. Samples were probed for adipolin (top panel) and only cell lysates were probed with GAPDH as a loading control (bottom panel).

No. NM_001014980.3, vector pcDNA3.1-C-(k)DYK) and verified by Sanger sequencing. Lysates and supernatants were collected after 72 hours from HEK293 cells transfected with the *C1QTNF12* overexpression plasmid and probed for adipolin expression by western blotting. Unconcentrated supernatant showed no detectable adipolin expression (data not shown); however, concentration of the supernatant by acetone precipitation confirmed the presence of adipolin (Figure 3.11).

To ascertain whether the 72-hour media conditioning process (i.e. 72-hour depletion/cell waste accumulation) would affect cellular function through factors unrelated to adipolin, cellular proliferation was compared between NDF cultured in control- and adipolin-conditioned media (72h-GFPCM, 72h-AdipCM). NDF cultured in undiluted 72h-GFPCM and 72h-AdipCM showed depressed proliferative capacity (Figure 3.12 A, B), leading to the conclusion that conditioning over this duration led to nutrient depletion. Comparing 24-hour conditioned media (24h-AdipCM and 24h-3.1CM) with various dilutions of standard complete growth media, proliferative capacity was unaffected (Figure 3.12C). Two independent collections of 24-hour conditioned media showed low but detectable levels of adipolin on enzyme-linked immunosorbant assay. Although the difference in adipolin concentration

between 24h-GFPCM and 24h-AdipCM was non-significant, there was a trend towards increased levels in the 24h-AdipCM (Figure 3.12D). Assuming this was a more physiological context for adipolin exposure, a decision was made to proceed with using 24h-AdipCM as the treatment for NDF to explore its potential transcriptional effects.

For the following RNA-sequencing experiments, the average mapping ratio with the reference genome was 95.47% and the average mapping ratio to genes was 87.33%; 17544 genes were identified. Differential expression analysis of NDF (n=2) over 48h in the presence of 24h-AdipCM showed that any transcriptional changes that resulted were subtle (Figure 3.13). Although sample segregation was present between treatment and control groups, this was eclipsed by the wide segregation between biological replicates (Figure 3.13A). As this was an exploratory experiment, analysis proceeded using DESeq2 to assess for differential expression signatures between AdipCM-treated and GFPCM-treated groups. Using an adjusted p-value threshold of 0.05, only 17 differentially expressed genes emerged when the log2 fold change threshold was set to 0 (Figure 3.13B, C); of these, Gene Ontology Enrichment Analysis yielded no statistically significant pathways.

Insufficient adipolin dosing by using conditioned media was a potential limitation of the insignificant transcriptomics findings. Attempts to purify recombinant adipolin from the conditioned media were unsuccessful (data not shown). As recombinant murine adipolin may cross react with its human counterpart (pairwise sequence alignment(Madeira et al., 2019) showed 75.8% sequence similarity between orthologous murine adipolin and human adipolin) and that its use on human aortic smooth muscle cells and human endothelial cells has been published(Enomoto et al., 2011; Ogawa et al., 2019), it was felt to be appropriate for testing on NDF. The recombinant murine adipolin was kindly gifted by Ouchi et al(Enomoto et al., 2011; Ogawa et al., 2019).

Being the only significant feature noted in the I1253 fibroblasts, when compared to WT NDF, TGFβ1-induced contractility was measured when exposed to the re-

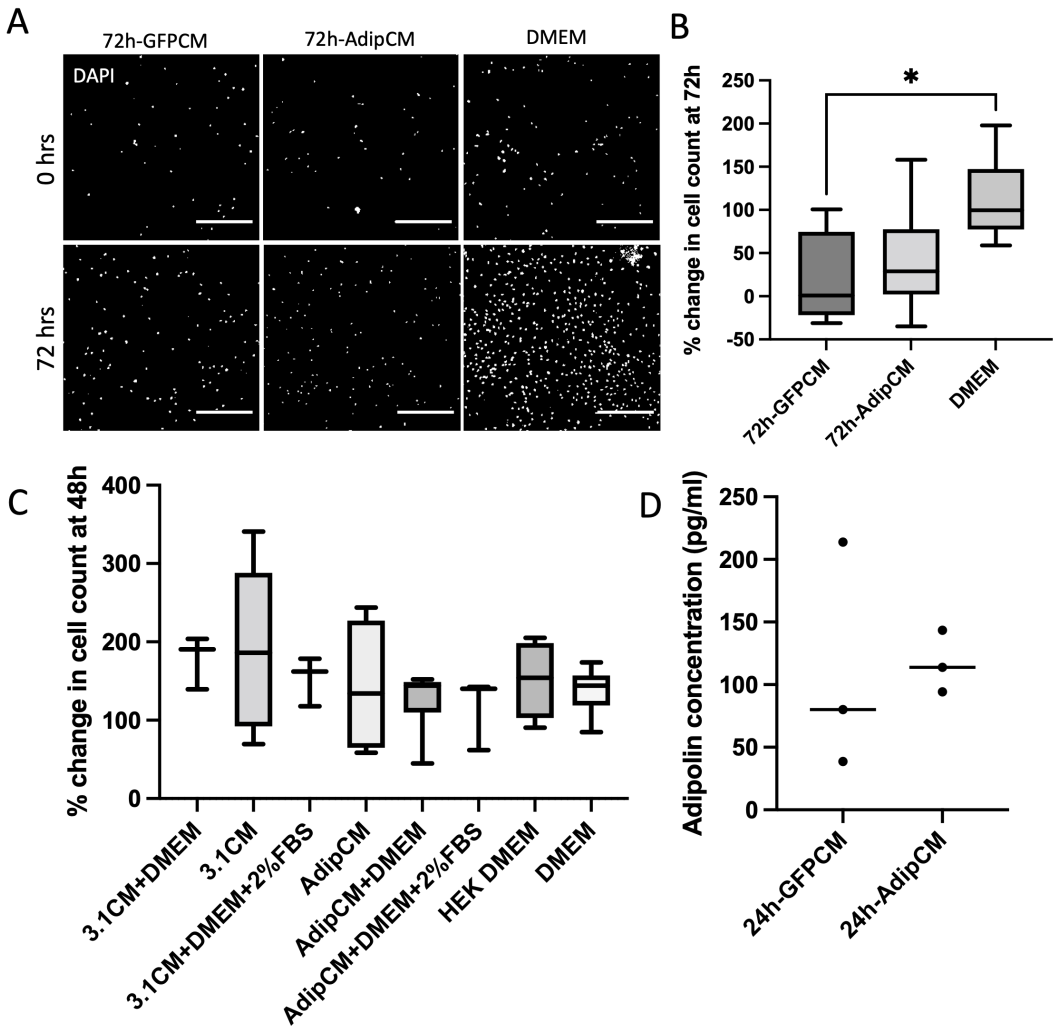

Figure 3.12: Fibroblasts treated with adipolin-conditioned media (AdipCM) show similar proliferative ability to control-conditioned media (green fluorescence protein-tagged, GFPCM or pcDNA3.1-tagged, 3.1CM). A. Example fluorescence microscopy images of NDF (n=2) treated with 72h-GFPCM, 72h-AdipCM and complete growth media supplemented with 10% FBS (DMEM), fixed at 6hr (top panels) and 72hr (bottom panels) growth and stained for DAPI. Scale bars: 400μm. B, C. Proliferation of NDF (n=2) cultured in various conditioned media dilutions quantified by counting change in DAPI-positive cells over time. Tile scans of the whole well were performed and stitched at 4x magnification. Box and whiskers plots of two biological replicates with minimum, maximum, 25th-75th percentile and median values indicated. * represents P <0.05 from 1-way ANOVA. D. Adipolin concentrations in 24h-GFPCM and 24h-AdipCM, measured using ELISA. Data represent three transfection samples from two biological replicates analysed in duplicates with the median value indicated.

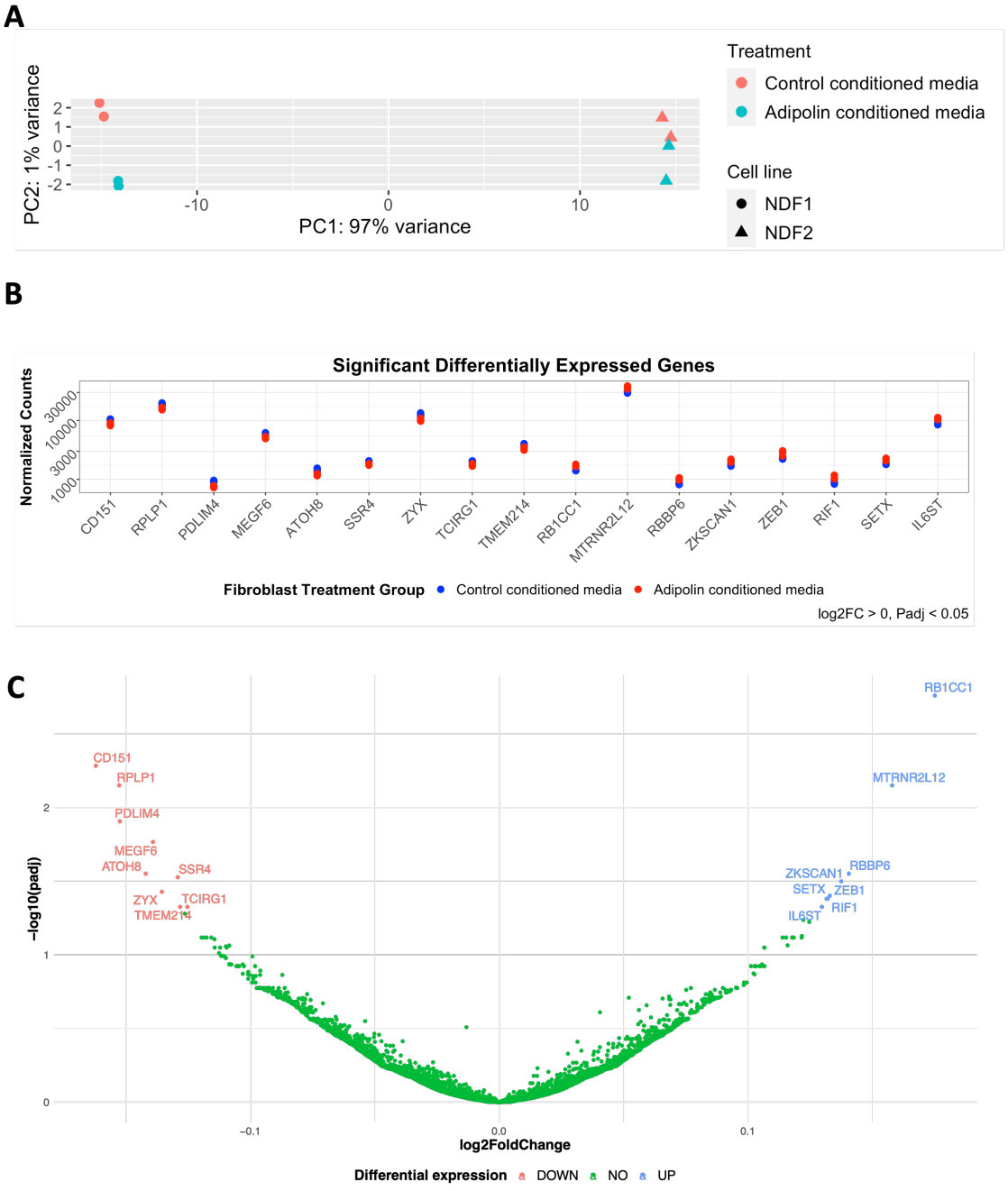

Figure 3.13: Fibroblasts treated with adipolin-conditioned media show subtle transcriptional changes. A. Principal component (PC) analysis plot of the transcriptome of normal fibroblasts (n=2) treated with and without adipolin-conditioned media, showing that ~97% of variance arises from the differences in biological samples and only 1% of variance arises from the conditioned media treatment. B, C. DESeq2 analysis using a liberal threshold for log2 fold change (log2FC) >0 and adjusted p-value (Padj) < 0.05 revealed few differentially expressed genes.

combinant murine adipolin. Using the published effective concentration of 300ng/ml, no significant difference in TGFβ1-induced contractility was detected (Figure 3.14).

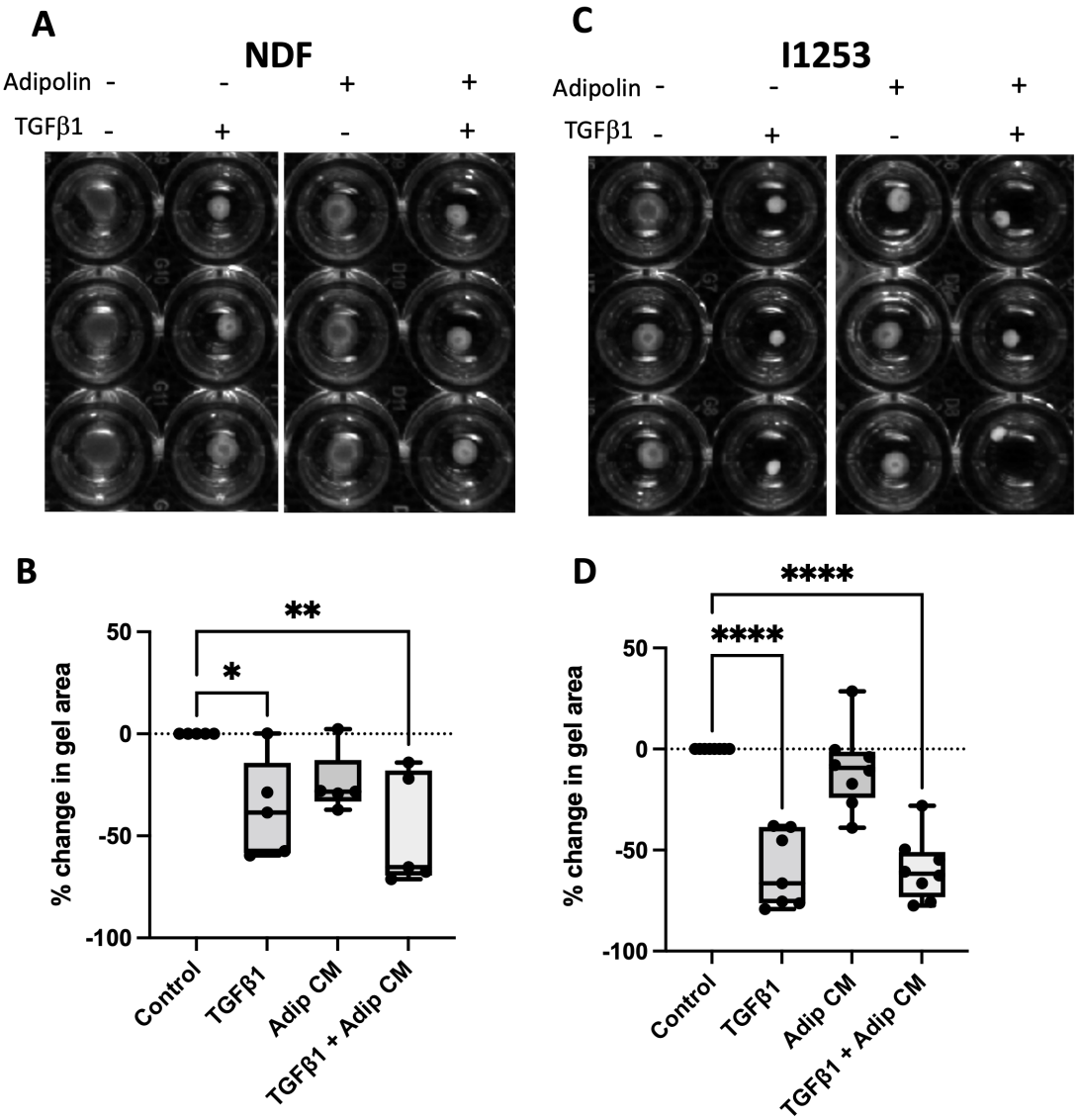

Figure 3.14: Treatment of dermal fibroblasts with recombinant mouse adipolin shows no significant changes to TGFβ1-induced contractility. A, C. Representative photographs of NDF- and I1253-embedded collagen gels treated with and without recombinant mouse adipolin (300ng/ml) and TGFβ1 (10ng/ml). B, D. Gel area quantified on ImageJ with baseline correction for control gels. Each point represents a gel from two experimental replicates.)

3.6 Phenotype of homozygous carriers of *C1QTNF12* LOF variant in the UKB

In light of the absence of detectable adipolin in dermal fibroblasts and the subtle phenotype observed in fibroblasts treated with adipolin (both adipolin-conditioned media and murine recombinant adipolin), the only evidence to support the hypothesis that this gene has an aetiological role in keloids is that its homozygous LOF variant segregated with this phenotype in a single small pedigree. An alternative approach was therefore sought for supporting evidence.

The UKB is a population based cohort study that recruited > 500000 participants (40-69 years old at recruitment) who attended one of 22 assessment centres across the UK between 2006 and 2010(Bycroft et al., 2018). A rich variety of health-related information is provided by the participants, including self-reported conditions, lifestyle indicators, genomic data and follow-up linked health records. Considering that the p.Asn43fs variant in *C1QTNF12* (rs571313759) has a reported minor allele frequency (MAF) of 0.0005 in Europeans and 0.0186 in Africans or African Americans according to gnomAD v3(Karczewski et al., 2020) it could be predicted that approximately two other individuals carrying this homozygous variant would be present in the UKB, which would allow for a qualitative assessment of the clinical implications.

The chromosomal region for *C1QTNF12* (chr1:1242453-1246722) was extracted from the UKB 200K exome release* and filtered for homozygous pLOF (defined as frameshift, stop-gained, splice-acceptor and splice-donor) variants. This resulted in four individuals; however, the genotype quality for their samples were low (Table 3.6). Three individuals, all of whom were homozygous carriers of rs571313759, had genotyping quality that indicated <20% error rate. Of these, two had linked electronic health records (EHR). A review of the summary diagnoses codes for both individuals showed no indications for primary fibrotic/scarring disease (Table 3.7).

*https://biobank.ndph.ox.ac.uk/ukb/label.cgi?id=170

Table 3.6: Participants carrying homozygous pLOF variants in *C1QTNF12* in the UKB and their genotyping quality.

Participant	Variant	Genotype Error Rate (%)
1	rs115005664	63.1
2	rs571313759	15.8
3	rs571313759	15.8
4	rs571313759	12.6

pLOF, predicted loss of function

Table 3.7: Diagnoses recorded for two individuals homozygous for the *C1QTNF12* pLOF variant, rs571313759, within the UKB.

Description	Group	n
Blood in stool	Digestive	1
Duodenal ulcer	Digestive	1
Duodenitis	Digestive	1
Haemetemesis	Digestive	1
Haemorrhage of gastrointestinal tract	Digestive	1
Other specified gastritis	Digestive	1
Calculus of kidney	Genitourinary	1
Renal colic	Genitourinary	1
Other anemias	Haematopoietic	1
Alcohol-related disorders	Mental Disorders	1
Alcoholic liver damage	Mental Disorders	1
Tobacco use disorder	Mental Disorders	1
Other headache syndromes	Neurological	1

pLOF, predicted loss of function; n indicates the number of individuals with the diagnosis in their electronic health records

3.7 PheWAS for heterozygous *C1QTNF12* LOF variant carriers

Novel disease associations have been reported particularly for rare protein coding variants and a case has been made for studying the additive effects of rare protein truncating variants(DeBoever et al., 2018). As such, to investigate whether pLOF variants in *C1QTNF12* contribute to risk for fibrotic/scarring traits, a PheWAS was performed for the 657 heterozygous carriers of pLOF *C1QTNF12* variants (all of whom had high genotyping quality i.e. genotype error rate <1%) using data extracted from the UKB 200K exomes (https://biobank.ndph.ox.ac.uk/ukb/label.cgi?id=170), with age, sex and ethnicity as covariates.

Of the 1522 phenotypes (phecodes, see Section 2.5.1) screened, there were no phenotypes that met phenome-wide significant associations (Figure 3.15). Of those that met nominal significance (p<0.05), the association with Dupuytren's disease was potentially of interest (odds ratio (OR) 2.04, p=0.027), although case numbers were small (there were ten individuals carrying pLOF variant(s) for *C1QTNF12* with a phecode diagnosis of Dupuytren's contracture compared with 585 without).

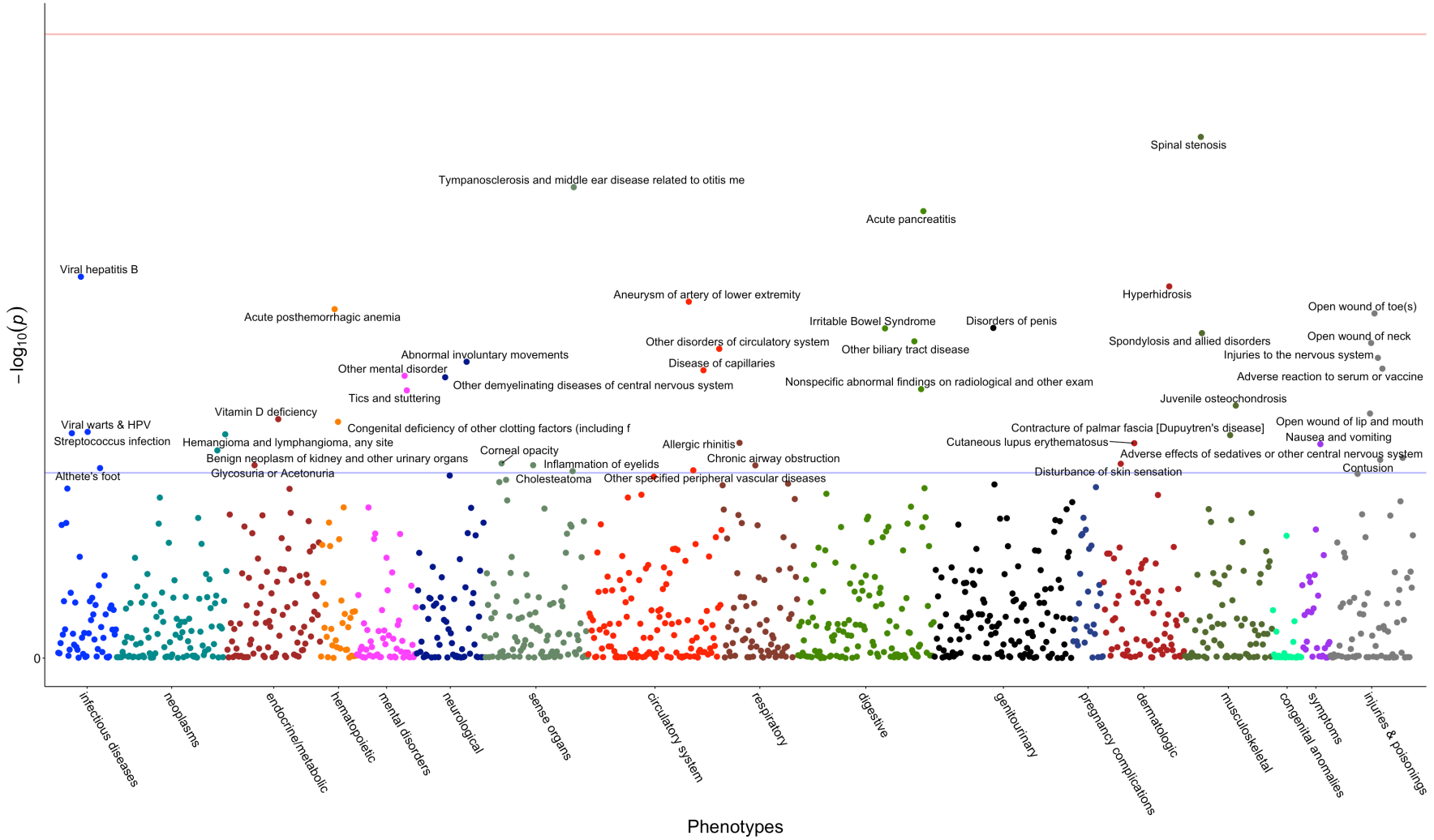

Figure 3.15: Manhattan plot showing multivariable logistic regression estimates for the effect of *C1QTNF12* pLOF heterozygosity on the risk of each phecode diagnosis, adjusting for age, sex and ethnicity. Dots represent phecodes and colours represent systemic categories. Orange line: Bonferroni-corrected significance threshold ($p<0.05/1518$, based on the number of phecodes tested). Blue line: Nominal significance threshold ($p<0.05$). pLOF, predicted loss of function.

3.8 Genotyping I1253 fibroblasts for single nucleotide polymorphisms relevant to keloids

In consideration that there may be factors other than the *C1QTNF12* variant responsible for or contributing to the phenotype of the pedigree, five SNPs that were previously identified to be associated with keloid were selected for Sanger sequencing, namely rs873549(Nakashima et al., 2010; Zhu et al., 2013), rs1511412(Nakashima et al., 2010; Zhu et al., 2013), rs940187(Nakashima et al., 2010; Zhu et al., 2013), rs8032158(Nakashima et al., 2010; Zhu et al., 2013), and rs192314256(Sakaue et al., 2021). The I1253 fibroblasts were found to be homozygous for the reference (i.e. non-risk) allele for all SNPs apart from rs8032158, located in intron 5 of the neuronal precursor cell-expressed developmentally downregulated 4 (*NEDD4*) gene, for which the fibroblasts were homozygous for the alternate (i.e. risk) allele. However, this variant did not segregate with the phenotype when Sanger sequencing was performed on the whole pedigree (Figure 3.16).

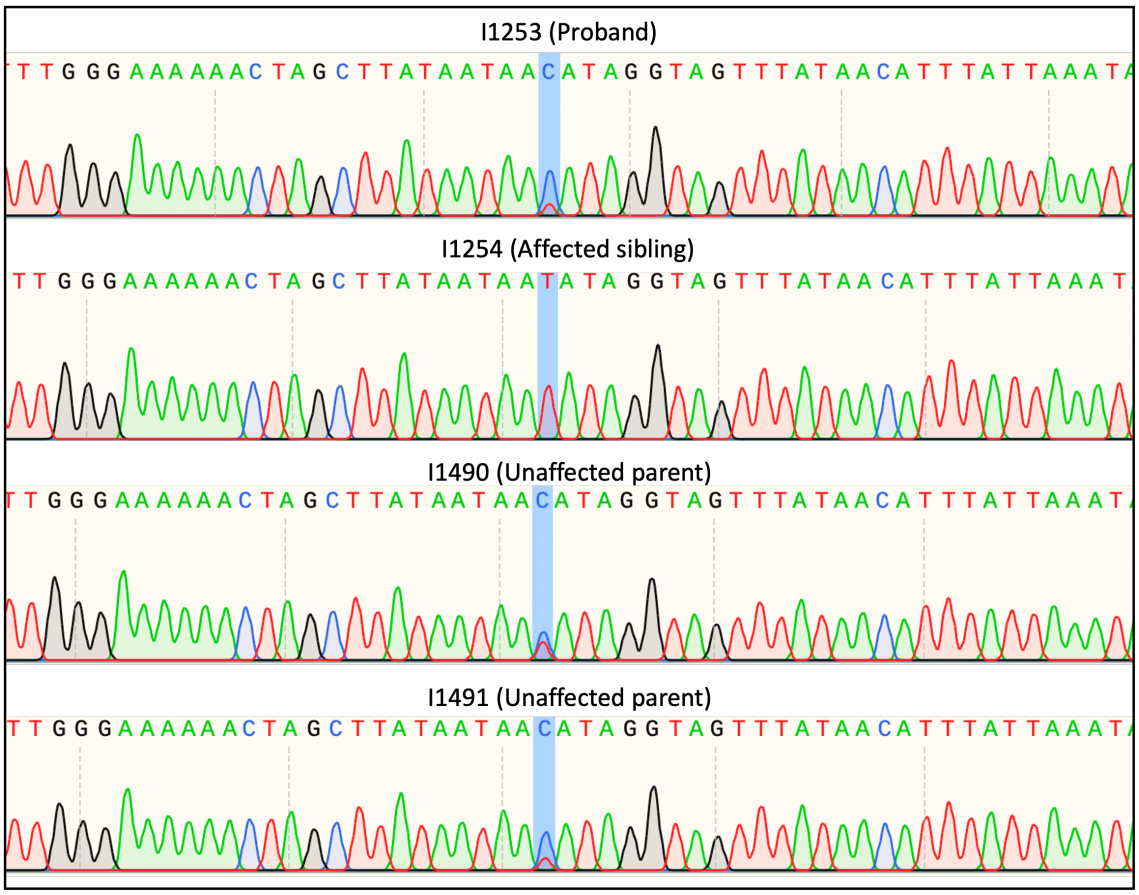

Figure 3.16: Sanger sequencing for rs8032158 (non-risk allele, T; risk allele, C; highlighted) did not show phenotype segregation (i.e. the risk allele was absent from the affected sibling and present in the unaffected parents).

3.9 Discussion

Experiments in this chapter started with studying the characteristics of I1253 fibroblasts (carrying the LOF variant in *C1QTNF12,* which encodes the insulin-sensitising adipokine, adipolin), derived from non-lesional skin of a patient suffering from excessive scarring. The hypothesis was that these fibroblasts are scar-prone as a result of the LOF variant. Although much has been published about KDFs, the characteristics of keloid-prone fibroblasts are less discussed. Single cell transcriptional profiling of mouse and human fibroblasts have indicated that CD26+ fibroblasts, which represent a large fraction of the human adult (non-papillary) dermis, may contribute to skin fibrosis and that differential expression of Wnt pathway, ECM and immunoregulatory genes may prime fibroblasts toward a pro- or anti-fibrotic role(Philippeos et al., 2018). Comparisons have also been made between oral mucosal fibroblasts (less prone to scarring) and skin fibroblasts (more prone to scarring) showing that skin fibroblasts are less proliferative and express a higher TGFβ1 signalling and cell-contractility gene signature(Mah et al., 2014). Consistent findings have been reported with comparisons between glabrous fibroblasts (less prone to scarring) and non-glabrous fibroblasts (more prone to scarring), showing that non-glabrous fibroblasts are less proliferative and accumulate more αSMA(Chipev and Simon, 2002).

The only significant finding in this chapter was that I1253 fibroblasts were 'myofibroblast-like', through the increased expression of αSMA and corresponding increased TGFβ1-induced contractility. αSMA expressing fibroblasts, typically deemed as differentiated myofibroblasts, are thought to be responsible for physiological wound contraction(Tomasek et al., 2002). Their persistance in a closed wound is typically considered to mediate pathological development of scar hypertrophy(Van De Water et al., 2013). It is likely that the contractile force generated by stress fibres is transmitted to the ECM leading to its straining and stiffening. ECM straining and stiffening activates TGFβ1(Froese et al., 2016; Klingberg et al., 2014), maintaining a feedback loop through which myofibroblast activity persists.

3. Adipolin and fibroblasts

Interestingly, there was no increase in type 1 collagen expression in the I1253 adipolin-mutant fibroblasts. Although generally assumed that collagen expression by KDFs is increased compared to normal fibroblasts(Li et al., 2011), it has been reported that fibroblasts isolated from keloids often synthesize normal amounts of collagen(Ala-Kokko et al., 1987; Barallobre-Barreiro et al., 2019).

The significance of I1253 fibroblasts being more 'myofibroblast-like' in culture may suggest they are profibrotic *in vivo*; however common cell culture conditions could be influencing this phenotype. Standard culture conditions are known to promote differentiation of primary fibroblasts into myofibroblasts(Baranyi et al., 2019). Furthermore, cells plated and passaged at a lower density are more likely to differentiate into myofibroblasts(Masur et al., 1996). Although effort was made to standardise culture conditions between patient and healthy control cells during the experiments, conditions prior to cryopreservation of the control cells were unknown.

Interrogation of the (autocrine) effect of adipolin in fibroblasts was hampered by the lack of its detectable autologous expression in the tissue types tested. Data presented here suggest that cultured dermal fibroblasts do not express adipolin at baseline or with attempted induction by glucose starvation, adipogenic differentiation and TGFβ1 treatment. Likewise, skin staining indicates absent adipolin expression although in this case there was no reliable positive control for the antibody tested (as was the case in data from the HPA, discussed in Section 1.3.2). Published literature has widely considered adipolin to be expressed in adipose tissue(Babapour et al., 2020; Enomoto et al., 2011; Sargolzaei et al., 2018; Tan et al., 2013). At study conception, it was hypothesized that adipolin may have relevance to fibroblasts, particularly as there is suggestion on the the public gene expression database, GTEx, that *C1QTNF12* is skin-enriched and that fibroblasts are present in adipose tissue(Gregoire, 2001). In hindsight, when compared with other adipokines (adiponectin, *ADIPOQ*; leptin, *LEP;* and resistin, *RETN)*(Rankinen et al., 2006) and skin-enriched genes (keratin 1, *KRT1*; periostin, *POTN*; and type 7 collagen, *COL7A1*)(Edqvist et al., 2014), its expression in both adipose tissue and skin is very low (Figure 3.17). Further examination of GTEx and ARCHS[4](Lachmann et

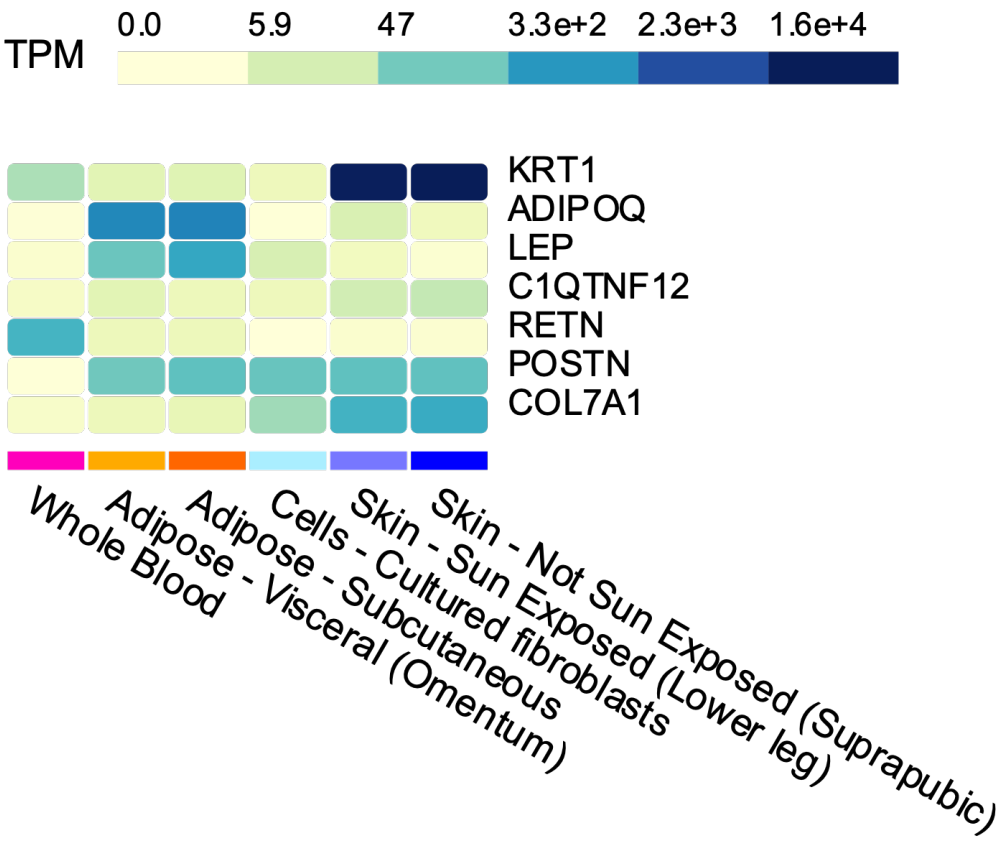

Figure 3.17: Comparison of *C1QTNF12* expression with selected adipokine- (*ADIPOQ*, *LEP*, *RETN*) and skin enriched- (*KRT1*, *POSTN*, *COL7A1*) genes. Data visualisation from the GTEx portal. TPM, transcripts per million.

al., 2018) databases suggested that adipolin gene expression is more prominent in the digestive and urogenital systems, but generally, constitutively low. As genes that are functionally significant tend to be more highly expressed in disease tissue(Feiglin et al., 2017), this supports the relevance of *C1QTNF12* for metabolic health, but less so for skin scarring.

There are notable limitations to studying adipolin expression. Data correlating transcriptomic and protein-level localisation in normal and excessive scars across different cell types is limited; signals for lowly expressed transcripts/proteins such as *C1QTNF12*/adipolin may have been diluted when analysing heterogeneous tissues/cells, as in the data sources examined. There are emerging technologies that allow spatially-resolved transcriptomic and protein profiling at subcellular levels (e.g. spatial transcriptomics(Marx, 2021) and imaging mass cytometry(Baharlou et

al., 2019)). Such information is important for scar research, considering their inter- and intra-lesional heterogeneity(Limandjaja et al., 2020b), and would be a valuable resource for future functional genetic studies (although caveats such as the need for effective and specific probes/antibodies, that may not be available for novel proteins, remain). Another key limitation is the absence of a temporal context in currently-available transcript/protein expression data to account for the highly dynamic processes in normal/excessive scars. Data on gene expression changes is lacking for wound healing, and even more so for excessive scars. Of publicly-available data, transcriptomic analysis of the skin of keloid-prone individuals at six weeks post wound healing did not highlight *C1QTNF12* as being differentially expressed(Onoufriadis et al., 2018). Similarly, changes in mouse skin transcriptome during wound healing did not highlight *C1QTNF12*(St. Laurent et al., 2017) although the duration studied (192 hours) corresponded with early-to-mid mammalian wound healing(Seifert et al., 2012), rather than the later scar-formation phase.

Nevertheless, efforts were made to investigate the effect of exogenous adipolin on cultured fibroblasts, in the event that fibroblasts were adipolin-responsive. Adipolin was detectable in concentrated conditioned media from adipolin-overexpressing HEK293 cells after 72 hours of conditioning. Downstream transcriptomics experiments utilised unconcentrated conditioned media after 24 hours of conditioning as 72-hour conditioning was found to hamper cellular proliferation. No significant differential gene expression was found through this approach. Here, two key limitations were the potential insufficient adipolin dosing in treatment conditions and the limited number of biological samples tested.

The focus of the *in vitro* study in this chapter has been on dermal fibroblasts but it is important to note that there are a myriad of other cell types involved in wound healing and scarring. As adipolin has previously been reported to be anti-inflammatory (reducing macrophage accumulation in adipose tissue(Enomoto et al., 2011)), it would be interesting to find out whether similar effects occur in skin, and whether an attenuated fibrotic response would ensue. For example, a published monocyte-keloid skin equivalent co-culture model(Limandjaja et al., 2019)

might be a starting point to study whether adipolin exposure influences macrophage phenotype switching (the M2 phenotype is thought to be contributory to keloid maintenance) as well as the increased contraction and αSMA staining reported by the authors in the keloid scar model.

Murine recombinant adipolin has previously been shown to protect against pathological vascular remodelling by reducing platelet-derived growth factor-BB-stimulated proliferation of human vascular smooth muscle cells through a TGFβ1-dependent pathway(Ogawa et al., 2019). Experiments in this chapter using murine recombinant adipolin at the published effective doses did not find evidence that it attenuated TGFβ1-induced fibroblast-mediated contraction of collagen gels. As protein availability was limited and adipolin pharmacokinetics have not been established, concurrent treatment of adipolin and TGFβ1 was employed. It is unknown whether pre-treatment of fibroblasts with adipolin may have stimulated anti-TGFβ1 signalling cascades when subsequently exposed to TGFβ1 (hence resulting in reduced contractility). However, although not statistically significant, it was interesting that exposure to adipolin alone appeared to promote rather than inhibit baseline (serum-induced) contraction.

Reviewing the phenotype of *C1qtnf12*-null mice in the Mouse Genome Initiative, the only reported abnormal phenotypes were enlarged/absent seminal vesicle and abnormal eye morphology (microphthalmia/anophthalmia)(Blake et al., 2020). In particular, there was no suggestion of abnormalities in the integumentary system. It is of interest that on metabolic challenges, phenotypes were subtle and sexually dimorphic, despite more promising *in vitro* reports(Tan et al., 2020). However, scar-modelling on these mice (e.g. by wounding or bleomycin treatment) may further the understanding of the relevance of this gene in fibrosis/excessive scars.

Considering the lack of strong functional evidence supporting the relevance of adipolin in the excessive scarring phenotype of the proband and the limitations of causal gene identification in a single small pedigree, data from the UKB was interrogated to identify carriers of homozygous pLOF variants of *C1QTNF12* and to describe their clinical manifestations. No supporting evidence was drawn from this

functional genomic analysis. This may have been a result of incomplete electronic health recording, incomplete phenotypic penetrance and/or the variable expressivity of phenotypic traits; however, it also calls into question whether the variant identified in this pedigree is truly causal of the excessive scarring phenotype in this pedigree.

The likelihood of rs571313759 in *C1QTNF12* leading to nonsense mediated decay remained an unproven hypothesis due to the inability to detect adipolin in the WT samples tested, as aforementioned. In light of the negative findings in this chapter, it would be sensible to consider the potential pathogenicity of other variants that cosegregated with the disease in the pedigree. The only other variant identified in the homozygous model was the missense variant in *SEMG1* (rs7263910), encoding semenogelin-1, the predominant protein in semen. Based on the tissue expression and published disease-associations of *SEMG1*(Pletscher-Frankild et al., 2015), this variant was felt to be less relevant. Reviewing the compound heterozygous variants shared by both siblings, a number of alternative variants may potentially be of relevance, namely *IGFN1* (rs138698246, rs368038036) and *NUDT19* (rs375826824, rs371803324), which were potentially pathogenic, based on prediction scores (Table 3.4). *IGFN1* encodes multiple splicing variants of Immunoglobulin- and Fibronectin-like domain containing proteins that are predominantly expressed in skeletal muscle(Li et al., 2017). There are reports of its protein variants being involved providing structural support to skeletal muscle sarcomere(Baker et al., 2010; Li et al., 2017) but the literature is sparse and the complexity of the locus makes it challenging to analyse. *NUDT19* encodes the Nudix hydrolase, Nudt19. Although largely understudied, Nudix hydrolases are involved in regulation of CoA, the obligate cofactor in numerous metabolic reactions(Shumar et al., 2018). Considering the hypothesised metabolic link in fibrosis(Ung et al., 2021), *NUDT19* may be a candidate for further study.

On the other hand, it is also worth considering the likely mode of inheritance of the disease. Although an autosomal recessive inheritance pattern was assumed in our pedigree, the most commonly reported inheritance pattern for familial keloids is autosomal dominant with incomplete penetrance and variable expression(Chen et

al., 2006a; Clark et al., 2009; Marneros et al., 2001; Ramakrishnan et al., 1974). As our pedigree only spanned two generations, the actual inheritance pattern could have been autosomal dominant with incomplete penetrance, opening up reams of alternative pathogenic genes. As our pedigree has since been lost to follow up, it was not possible to perform a more extensive pedigree review alongside more detailed phenotyping. This would be an important future step, should the opportunity for study arise, as larger pedigrees would increase power to enable gene mapping and rare variant identifcation(Wijsman, 2012). It would also be of interest to revisit previously reported keloid pedigrees that are likely to have enlarged over time.

Overall, findings in this chapter do not provide functional evidence to support the hypothesis that *C1QTNF12* LOF variant leads to a profibrotic phenotype. Caution has been raised on the interpretation of rare homozygous LOF genotypes identified by exome or genome sequencing, as although the effect on protein function is greatest, most of the proteins affected are LOF-tolerant(Narasimhan et al., 2016a). *C1QTNF12* has not yet been reported in systematic surveys with phenotypic analyses of homozygous LOF variants(Lim et al., 2014; MacArthur et al., 2012; Narasimhan et al., 2016a; Saleheen et al., 2017). This chapter illustrates the need for meticulous and systematic phenotyping of pedigrees subjected to DNA sequencing, in tying together their clinical, epidemiological and molecular data, to identify novel biological functions for the highlighted genes.

3.9.1 Limitations

A key limitation of the *in vitro* work is having n=1 of patient fibroblasts, resulting in a lack of power to detect gene expression differences that might explain the 'myofibroblast' phenotype observed in the patient cells. There is increasing discussion surrounding the heterogeneity of fibroblasts, depending on tissue status, regional features, microenvironment and cellular state(Shaw and Rognoni, 2020). Dissecting differences in fibroblasts between disease and healthy states whilst capturing their physiological complexity is challenging, ideally requiring large sample sizes that are adequately stratified, tested using a model that recapitulates an *in vivo* state.

3. Adipolin and fibroblasts

The lack of a known receptor for adipolin also limited the approaches undertaken to investigate its effect. Its future identification and a knowledge of its tissue distribution would aid the study of adipolin-responsive tissues and cells. Studying the adipolin ligand may provide supporting evidence to its function. As well as *in vitro* experiments, this includes the bioinformatic study of phenotypic associations of its genetic variants at a population level, expanding on the approach undertaken in this chapter.

Interrogating the phenotype of homozygous *C1QTNF12*-LOF carriers was limited by having only two additional individuals with linked-EHR. This is expected, due to the rarity of this variant. Gene-based variant burden testing is a statistically robust method to screen for pathological phenotypes associated with rare variants of this gene - a comprehensive database for this was recently made publicly-available(Karczewski et al., 2022). Through this resource, analysis of *C1QTNF12* rare variant burden testing was unremarkable; however, the data used did not include primary health records and may have excluded a significant number of relevant diagnoses, including keloids.

4

Determining comorbidity links with excessive scarring

Contents

This chapter presents results from an epidemiological assessment of diseases that are interrelated with keloids and/or hypertrophic scars (henceforth "excessive scarring") in the UK Biobank (UKB), a multi-centre population-based longitudinal observational study of >500000 participants including 972 with a diagnosis of excessive scarring. A candidate approach was first undertaken to validate previously reported disease associations, followed by a discovery approach, using a PheWAS.

A PheWAS is an unbiased exploration of associations between independent variables (e.g. genetic variants/biological measurement/clinical phenotype) and the entire range of clinical phenotypes (the phenome)(Hebbring, 2014). These phenotypes are designated as phecodes, curated hierarchical groups of clinical codes

for research, covering broad organ systems/categories(Bastarache, 2021). Although primarily used for billing and statistical reporting purposes, clinical coding systems are fundamental to the comprehensive electronic health information accessible for research. However, classification systems vary between settings; in the UKB, secondary care health records are coded using the ICD system (ICD9 and ICD10), whereas primary care data are coded using Read version 2 (Read2) and Read version 3 (Read3) terminologies(Sudlow et al., 2015). Translation of clinical codes between the different classifications is required to enable their integration for research. In this study, clinical codes retrieved from the UKB were mapped to their corresponding (ICD10-based) phecodes before multivariable association testing.

4.1 Study cohort

230078 UKB participants for whom linked GP data were available were analysed (Figure 4.1). 972 participants had a record of excessive scarring (740 with a diagnostic code specific for keloid, 110 specific for hypertrophic scar, 177 for either keloid or hypertrophic scar; Figure 4.2) compared to 229106 participants without. Table 4.3 shows the baseline characteristics for the excessive scarring and comparator groups. In the excessive scar-affected group, there was a higher proportion of females than in the unaffected group (65% versus 55%) and a lower proprtion of participants with self-reported white ethnicity (86% versus 95%).

There were 3403 individuals (1.2% of the study cohort) with missing data for at least one of age, sex, ethnicity, Townsend Deprivation Index, body mass index (BMI) or smoking status. Participants with missing data were more likely to be male and of non-white ethnicity, with higher Townsend Deprivation Index and BMI (Table 4.1). Comparing participants with excessive scarring diagnoses versus those without, there were no significant differences in the proportion of missing data for these variables (Table 4.2). Nevertheless, for testing associations, the full cohort was analysed and multiple imputation was used to account for missing data.

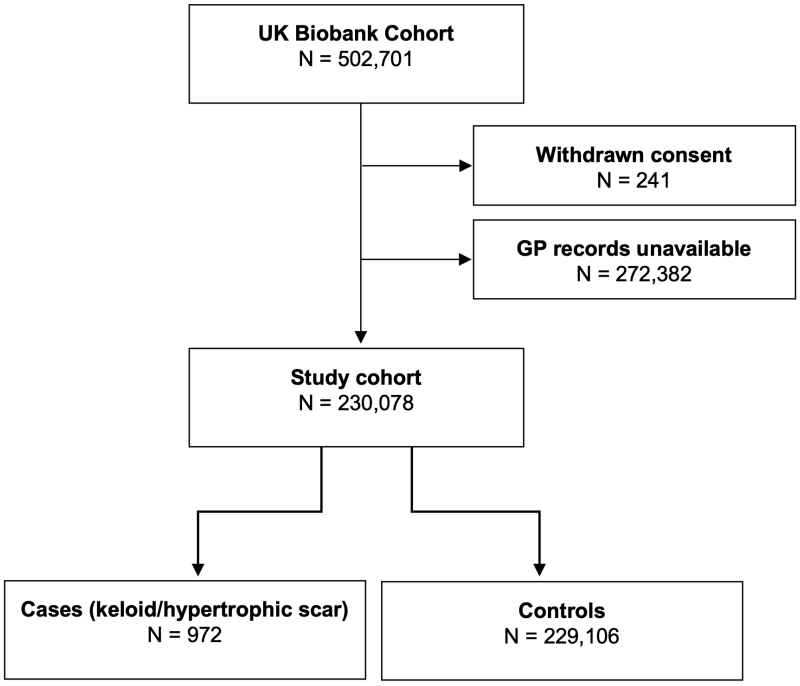

Figure 4.1: Flowchart of participants included in study cohort. Participants without linked primary care data were excluded.

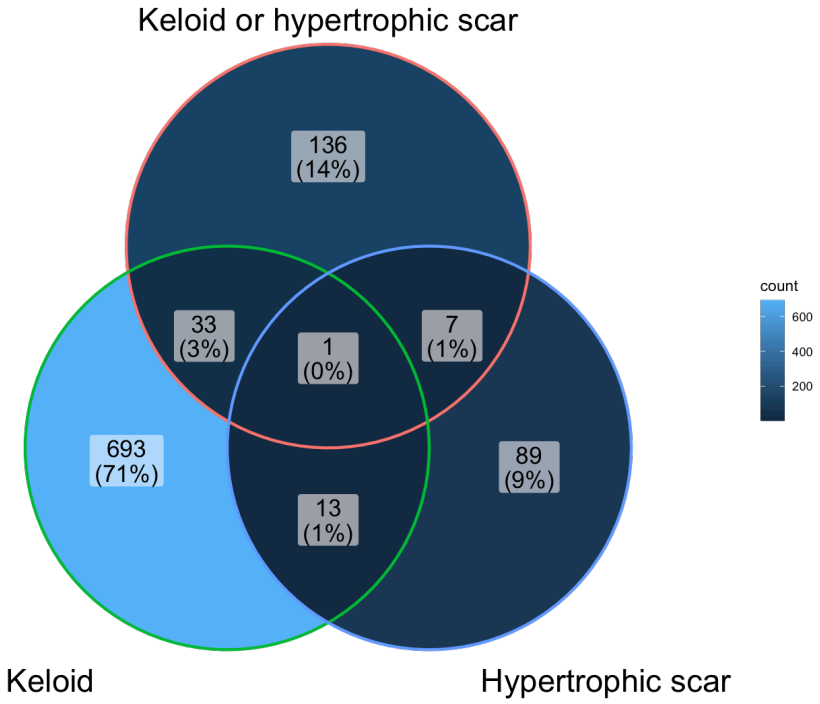

Figure 4.2: Venn diagram of patients with keloid and/or hypertrophic scar diagnosis codes in their linked electronic health records.

Table 4.1: Summary of individuals with and without missing data.

Characteristic	With missing data, N = 3,403	Without missing data, N = 226,675	p-value
Age	65 (57, 71)	66 (58, 71)	0.4
Sex			**<0.001**
Female	1,668 (49%)	124,103 (55%)	
Male	1,735 (51%)	102,572 (45%)	
Ethnic background			**<0.001**
White	1,958 (84%)	216,373 (95%)	
Asian or Asian British	205 (8.8%)	4,502 (2.0%)	
Black or Black British	81 (3.5%)	2,428 (1.1%)	
Chinese	14 (0.6%)	586 (0.3%)	
Mixed	10 (0.4%)	1,157 (0.5%)	
Other ethnic group	57 (2.5%)	1,629 (0.7%)	
Townsend Deprivation Index	-0.62 (-2.92, 2.63)	-2.16 (-3.65, 0.45)	**<0.001**
Ever smoked	1,292 (58%)	134,654 (59%)	0.3
Body mass index	27.4 (24.7, 31.0)	26.8 (24.2, 30.0)	**<0.001**
Keloid or Hypertrophic Scar			0.7
Control	3,387 (100%)	225,719 (100%)	
Case	16 (0.5%)	956 (0.4%)	
Hypertension	1,430 (42%)	77,604 (34%)	**<0.001**
Leiomyoma of uterus	196 (5.8%)	13,935 (6.1%)	0.3
Osteoporosis	243 (7.1%)	13,071 (5.8%)	**<0.001**

Continuous variables are summarised as median (interquartile range) and compared using Wilcoxon rank sum test. Categorical variables as number (%) and compared using Pearson's Chi-squared test. Statistical significance is declared at $P < 0.05$ (bold).

Table 4.2: Missing data for keloid or hypertrophic scar cases versus control individuals.

Characteristic	Control, N = 229,106[1]	Case, N = 972[1]	p-value[2]
Age			
Sex			
Townsend Deprivation Index			>0.9
Missing	342 (0.1%)	1 (0.1%)	
Ethnic background			0.8
Missing	1,073 (0.5%)	5 (0.5%)	
Ever smoked			0.6
Missing	1,187 (0.5%)	4 (0.4%)	
Body mass index			>0.9
Missing	1,352 (0.6%)	6 (0.6%)	

[1]n (%)

[2]Fisher's exact test; Pearson's Chi-squared test

Statistical significance is declared at $P < 0.05$.

Table 4.3: Baseline characteristics for the excessive scarring and comparator groups.

	Study Cohort	**Keloid or Hypertrophic Scar Cohort**		
Characteristic	**N = 230,078**	**Case, N = 972**	**Control, N = 229,106**	**p-value**
Age	64 (8)	63 (8)	64 (8)	**<0.001**
Sex				**<0.001**
Female	125,771 (55%)	633 (65%)	125,138 (55%)	
Male	104,307 (45%)	339 (35%)	103,968 (45%)	
Ethnic Background				**<0.001**
Asian or Asian British	4,707 (2.1%)	53 (5.5%)	4,654 (2.0%)	
Black or Black British	2,509 (1.1%)	61 (6.3%)	2,448 (1.1%)	
Chinese	600 (0.3%)	7 (0.7%)	593 (0.3%)	
Mixed	1,167 (0.5%)	7 (0.7%)	1,160 (0.5%)	
Other ethnic group	1,686 (0.7%)	10 (1.0%)	1,676 (0.7%)	
White	218,331 (95%)	829 (86%)	217,502 (95%)	
Townsend Deprivation Index	-1.33 (3.03)	-1.29 (3.12)	-1.33 (3.03)	0.7
Ever Smoked	135,946 (59%)	526 (54%)	135,420 (59%)	**0.001**
Body Mass Index	27.5 (4.8)	27.9 (5.2)	27.5 (4.8)	0.051
Hypertension	79,034 (34%)	362 (37%)	78,672 (34%)	0.062
Leiomyoma of uterus[a]	14,131 (6.1%)	92 (9.5%)	14,039 (6.1%)	**<0.001**
Vitamin D deficiency	5,547 (2.4%)	50 (5.1%)	5,497 (2.4%)	**<0.001**
Eczema, atopic	13,501 (5.9%)	99 (10%)	13,402 (5.8%)	**<0.001**
Leiomyoma of uterus	14,080 (11%)	92 (15%)	13,988 (11%)	**0.009**

Continuous variables are summarised as mean (standard deviation) and compared using Welch's t-test. Categorical variables as number (%) and compared using Pearson's Chi-squared test. Statistical significance is declared at $P < 0.05$ (bold); [a]Only female participants (N total = 125,771, N affected = 633, N unaffected = 125,138) considered for uterine leiomyoma.

4.2 Previously-studied associations with excessive scarring

The systematic search for excessive scarring disease associations resulted in 708 references in independent English language studies that were independently reviewed for relevant studies, excluding case reports. From the 21 remaining references (Table 4.4), disease associations selected for analysis in this study were those that were studied in two or more independent reports, namely hypertension, uterine leiomyoma, vitamin D deficiency and atopic eczema.

Table 4.4: Summary of systematic search for published keloid and hypertrophic scar disease associations.

PMID	Title	Study type	Participants	Disease Association	Disease risk
26677295	The Association between Atopic Disorders and Keloids: A Case-control Study.	Cohort	European	Atopic asthma	Increase
34880399	The increased prevalence of keloids in atopic dermatitis patients with allergic comorbidities: a nationwide retrospective cohort study.	Cohort	East Asian	Atopic dermatitis	Increase
30021755	Keloid risk in patients with atopic dermatitis: a nationwide retrospective cohort study in Taiwan.	Cohort	East Asian	Atopic dermatitis	Increase
28497468	A case-control study analyzing the association of keloids with hypertension and obesity.	Case control	African American	Hypertension	Increase

PMID	Title	Study type	Participants	Disease Association	Disease risk
26517298	Association of keloids with systemic medical conditions: a retrospective analysis.	Case control	African American	Hypertension	Increase
26517298	Association of keloids with systemic medical conditions: a retrospective analysis.	Case control	African American	Obesity	Increase
25728259	Hypertension: a systemic key to understanding local keloid severity.	Cohort	East Asian	Hypertension	Nil significant
12459533	A controlled cohort study examining the onset of hypertension in black patients with keloids.	Case control	Blacks (not specified)	Hypertension	Increase
8618497	Keloid associated with hypertension.	Case control	African Americans and whites	Hypertension	Increase
33541932	Migraines and keloids: a 15-year Taiwan claim database analysis.	Case control	East Asian	Migraine	Increase
33413286	The association between keloid and osteoporosis: real-world evidence.	Case control	East Asian	Osteoporosis	Increase
34697275	Is Hypertrophic or Keloid Wound Scar a Risk Factor for Stricture at Esophagogastric Anastomosis Site after Esophageal Cancer Operation?.	Cohort	East Asian	Post-procedural anastamotic stricture	Increase

PMID	Title	Study type	Participants	Disease Association	Disease risk
33403515	Arthrofibrosis after total knee arthroplasty: patients with keloids at risk.	Case control	Not specified	Post-procedural arthrofibrosis	Increase
33931723	Risk of cancer development in patients with keloids.	Case control	East Asian	Skin cancer	Increase
33931723	Risk of cancer development in patients with keloids.	Case control	East Asian	Pancreatic cancer (females only)	Increase
25081927	Keloid incidence in Asian people and its comorbidity with other fibrosis-related diseases: a nationwide population-based study.	Case control	East Asian	Uterine leiomyoma	Increase
24386410	Keloids and ultrasound detected fibroids in young African American women.	Cohort	African American	Uterine leiomyoma	Nil significant
33476473	Does vitamin D deficiency predispose to keloids via dysregulation of koebnerisin (S100A15)? A case-control study.	Case control	Asian (Egyptian)	Vitamin D deficiency	Increase
23867793	The TaqI gene polymorphisms of VDR and the circulating 1,25-dihydroxyvitamin D levels confer the risk for the keloid scarring in Chinese cohorts.	Cohort	East Asian	Vitamin D deficiency	Increase

PMID	Title	Study type	Participants	Disease Association	Disease risk
30204734	Effect of vitamin D deficiency on hypertrophic scarring.	Cohort	Asian (Turkish)	Vitamin D deficiency	Increase
30806461	The association between postburn vitamin D deficiency and biomechanical properties of hypertrophic scars	Cohort	East Asian	Vitamin D deficiency	Increase

All previously-studied comorbidities were more prevalent for individuals with excessive scarring. Vitamin D deficiency and atopic eczema were around twice as common in the excessive scarring cohort: 5.1% versus 2.4% (vitamin D deficiency) and 10% versus 5.8% (atopic eczema). The differences in uterine leiomyoma and hypertension prevalence were smaller: 15% versus 11% (leiomyoma) and 37% versus 34% (hypertension).

Each comorbidity was analysed as a disease outcome using two multivariable logistic regression models: minimally-adjusted (adjusting for age, sex and ethnicity, except for uterine leiomyoma which was tested within females only adjusting for age and ethnicity) and fully-adjusted (with additional potential confounders for each comorbidity; Table 4.5). Statistically-significant associations with excessive scarring were observed for hypertension and atopic eczema in the minimally-adjusted models, while the association with vitamin D deficiency fell short of Bonferroni-corrected significance (OR 1.42, p=0.022). In fully-adjusted models, only the association with atopic eczema (OR 1.68, $p<0.001$) was significant. Despite a positive effect-size estimate, there was no significant association between excessive scarring and uterine leiomyoma.

Table 4.5: Associations between excessive scarring and selected comorbidities.

	Minimal Model			Full Model		
Outcome	**OR[1]**	**95% CI[1]**	**p-value**	**OR[1]**	**95% CI[1]**	**p-value**
Hypertension	1.24	1.08, 1.43	**0.002**	1.11	0.96, 1.30	0.2
Uterine leiomyoma	1.20	0.96, 1.51	0.11	1.19	0.95, 1.49	0.13
Vitamin D deficiency	1.42	1.05, 1.93	0.022	1.47	1.09, 1.99	0.013
Atopic eczema	1.78	1.44, 2.19	**<0.001**	1.68	1.36, 2.07	**<0.001**

[1]OR = Odds Ratio, CI = Confidence Interval

Minimal model adjusts for age, sex (except uterine leiomyoma) and ethnicity for the entire imputed cohort. Full model adjusts for age, sex (except uterine leiomyoma), ethnicity and additional confounders for the entire imputed cohort (see Methods). Models for uterine leiomyoma are restricted to females only without adjusting for sex.

Statistical significance is declared at $P < 0.05/4 = 0.0125$ (bold)

Association testing was then performed within subgroups of participants defined by self-reported ethnicity. Results for white (829 excessive scar affected, 217502 unaffected), black (61 affected, 2448 unaffected) and Asian (53 affected, 4654 unaffected) participants are summarized in Table 4.6. With the exception of atopic eczema, the analysis revealed a divergence between ethnic groups in the prevalence of each comorbidity and its association with excessive scarring (Table 4.7). The associations with hypertension and uterine leiomyoma were nominally significant in black participants [OR 2.05, p=0.019 (hypertension) and OR 1.93, p=0.05 (uterine leiomyoma)] and not significant in white or Asian participants. Vitamin D deficiency was only significantly associated with excessive scarring in Asian participants (OR 2.24, p=0.006). For atopic eczema, the association with excessive scarring was highly

significant in white participants (OR 1.69, p<0.001), nominally significant in Asian participants (OR 2.17, p=0.048), and although not statistically significant in black participants, exhibited a similar trend (OR 1.89, p=0.13).

Table 4.6: Summary of comorbidities in cases and controls by ethnic groups.

	White participants			Black participants			Asian participants		
Characteristic	**Control, N = 217,502[1]**	**Case, N = 829[1]**	**p-value[2]**	**Control, N = 2,448[1]**	**Case, N = 61[1]**	**p-value[2]**	**Control, N = 4,654[1]**	**Case, N = 53[1]**	**p-value[2]**
Hypertension	74,175 (34%)	290 (35%)	0.6	1,056 (43%)	38 (62%)	**0.004**	1,995 (43%)	23 (43%)	>0.9
Vitamin D deficiency	3,866 (1.8%)	20 (2.4%)	0.2	263 (11%)	6 (9.8%)	>0.9	1,041 (22%)	20 (38%)	**0.013**
Eczema, atopic	12,621 (5.8%)	83 (10%)	**<0.001**	144 (5.9%)	7 (11%)	0.12	323 (6.9%)	8 (15%)	**0.042**
Uterine leiomyoma	12,843 (11%)	63 (12%)	0.6	458 (33%)	18 (47%)	0.083	291 (13%)	4 (14%)	>0.9

[1]n (%)

[2]Pearson's Chi-squared test

Statistical significance is declared at $P < 0.05$.

Table 4.7: Associations between pathological scars and selected comorbidities within three main UKB self-reported ethnic groups.

	White partcipants			Black participants			Asian participants		
Outcome	**OR[1]**	**95% CI[1]**	**p-value**	**OR[1]**	**95% CI[1]**	**p-value**	**OR[1]**	**95% CI[1]**	**p-value**
Hypertension	1.08	0.92, 1.28	0.3	2.05	1.13, 3.72	0.019	0.95	0.48, 1.86	0.9
Uterine leiomyoma*	1.07	0.82, 1.39	0.6	1.93	1.00, 3.71	0.050	1.04	0.36, 3.01	>0.9
Vitamin D deficiency	1.32	0.85, 2.07	0.2	0.88	0.37, 2.08	0.8	2.24	1.26, 3.97	**0.006**
Atopic eczema	1.68	1.34, 2.12	**<0.001**	1.89	0.83, 4.28	0.13	2.17	1.01, 4.67	0.048

[1]OR = Odds Ratio, CI = Confidence Interval

Statistical significance is declared at $P < 0.05/4 = 0.0125$ (bold)

*Model for uterine leiomyoma is restricted to females only without adjusting for sex.

4.3 Discovery analysis

1518 phecodes across 17 disease groups were screened, identifying 110 diseases significantly enriched ($p<0.05/1518=3.3\times10^{-05}$ based on 1518 phecodes) among participants with excessive scarring (Figure 4.3, Table 4.8, Table 4.9).

There was strongest evidence of association for several dermatological diseases, most prominently sebaceous cyst (OR 2.56, $p=9.45x10^{-30}$), non-epithelial skin cancer(OR 2.89, $p=2.03x10^{-25}$) and the umbrella phenotype "diseases of hair/hair follicles" (OR 2.3, $p=1.50x10^{-22}$), as well as infections of skin/subcutaneous tissue, seborrheic keratosis, actinic keratosis, acne, and notably, atopic/contact dermatitis, all with OR>1.9 and $p<1.0x10^{-11}$. Similarly strong evidence was observed for pain-related symptoms, particularly for joint pain (OR 1.84, $p=1.87x10^{-20}$) but also back pain, cervicalgia, enthesopathies and mastodynia, all with OR>1.6 and $P< 1.0x10^{-12}$. Significant associations with the largest effect sizes (when limited to those with more than 10 cases with excessive scarring) were abnormal weight gain (OR 3.97, $p=9.32x10^{-8}$), heart valve replacement (OR 3.9, $p=6.65x10^{-7}$) and hypertrophy of breast (OR 3.82, $p=9.53x10^{-10}$).

Associations were identified with hypertension (OR 1.26, $p=2.05x10^{-3}$) and vitamin D deficiency (OR 1.47, $p=1.34x10^{-2}$) as expected, but these did not meet Bonferroni-corrected significance.

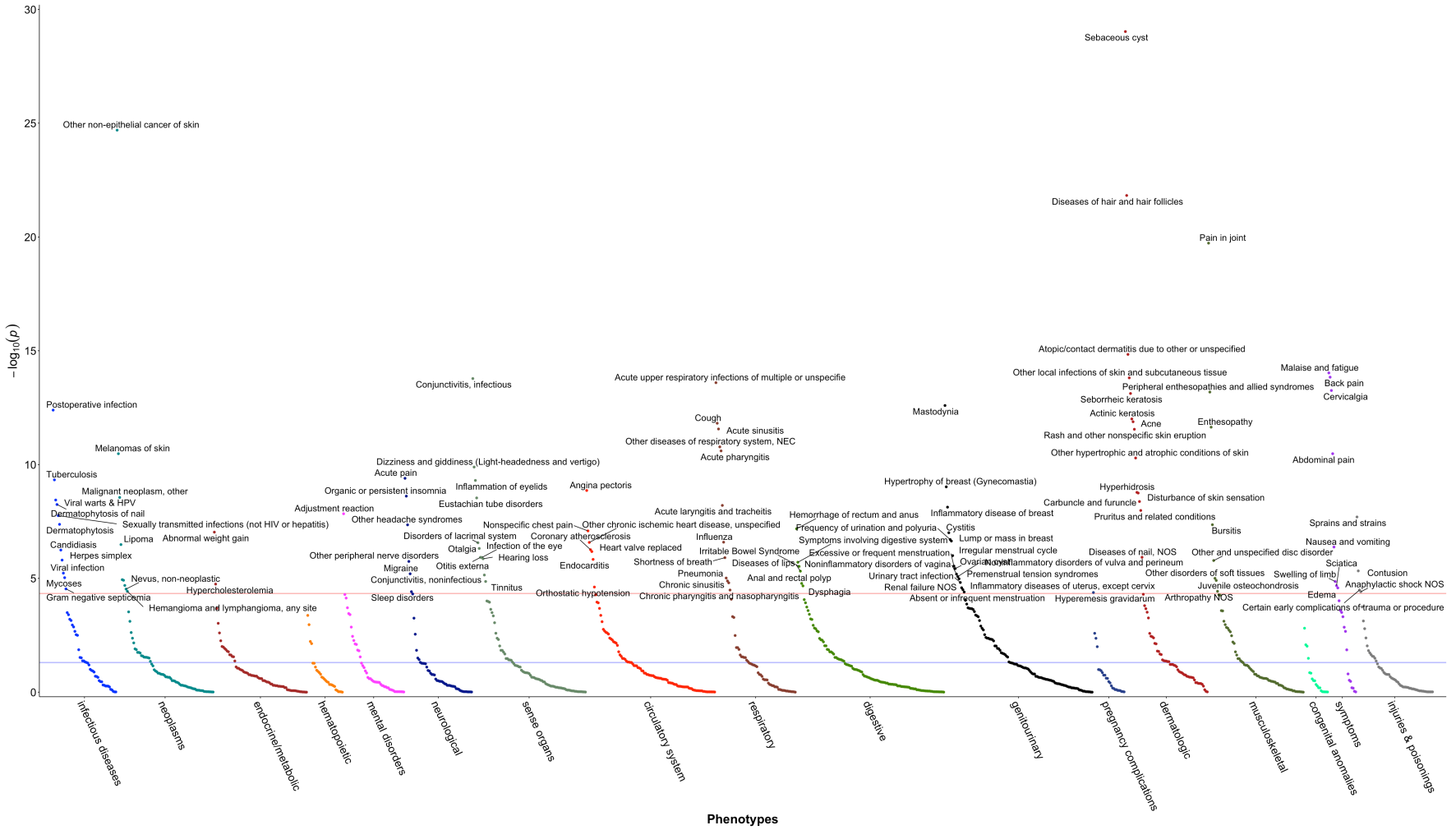

Figure 4.3: Manhattan plot showing multivariable logistic regression estimates for the effect of excessive scarring status on the risk of each phecode diagnosis, adjusting for age, sex, ethnicity, smoking status, body mass index and Townsend Deprivation Index. Dots represent phecodes and colours represent systemic categories. Statistical significance was set at $P < 0.015/1518$ (based on the number of phecodes tested), indicated by the red horizontal line.

Table 4.8: Phecodes significantly associated with excessive scarring status.

Group	Phecode description
Genitourinary (15)	**Mastodynia**; Hypertrophy of breast (Gynecomastia); Inflammatory disease of breast; Cystitis; Frequency of urination and polyuria; Lump or mass in breast; Irregular menstrual cycle; Excessive or frequent menstruation; Ovarian cyst; Noninflammatory disorders of vagina; Noninflammatory disorders of vulva and perineum; Premenstrual tension syndromes; Urinary tract infection; Renal failure NOS; Inflammatory diseases of uterus, except cervix
Dermatologic (14)	**Sebaceous cyst**; **Diseases of hair and hair follicles**; **Atopic/contact dermatitis due to other or unspecified**; **Other local infections of skin and subcutaneous tissue**; **Seborrheic keratosis**; **Actinic keratosis**; **Acne**; Rash and other nonspecific skin eruption; Other hypertrophic and atrophic conditions of skin; Hyperhidrosis; Disturbance of skin sensation; Carbuncle and furuncle; Pruritus and related conditions; Diseases of nail, NOS
Respiratory (12)	**Acute upper respiratory infections of multiple or unspecified sites**; **Cough**; **Acute sinusitis**; Other diseases of respiratory system, NEC; Acute pharyngitis; Acute laryngitis and tracheitis; Influenza; Shortness of breath; Pneumonia; Pneumonia due to fungus (mycoses); Chronic sinusitis; Chronic pharyngitis and nasopharyngitis
Infectious diseases (11)	**Postoperative infection**; Tuberculosis; Viral warts & HPV; Dermatophytosis of nail; Sexually transmitted infections (not HIV or hepatitis); Dermatophytosis; Candidiasis; Herpes simplex; Viral infection; Mycoses; Gram negative septicemia
Sense organs (11)	**Conjunctivitis, infectious**; Dizziness and giddiness (Light-headedness and vertigo); Inflammation of eyelids; Eustachian tube disorders; Disorders of lacrimal system; Otalgia; Otitis externa; Infection of the eye; Hearing loss; Conjunctivitis, noninfectious; Tinnitus
Neoplasms (8)	**Other non-epithelial cancer of skin**; Melanomas of skin; Malignant neoplasm, other; Lipoma; Other benign neoplasm of connective and other soft tissue; Benign neoplasm of lip, oral cavity, and pharynx; Malignant neoplasm of female breast; Nevus, non-neoplastic
Symptoms (8)	**Malaise and fatigue**; **Back pain**; **Cervicalgia**; Abdominal pain; Nausea and vomiting; Swelling of limb; Sciatica; Edema
Circulatory system (7)	Angina pectoris; Nonspecific chest pain; Other chronic ischemic heart disease, unspecified; Coronary atherosclerosis; Heart valve replaced; Endocarditis; Orthostatic hypotension

Significant associations with excessive scarring ($p < 0.05/1518 = 3.3\times10\text{-}05$), grouped by category (number of phecodes per category shown in brackets). The 20 most significantly associated phecodes are in bold.

Group	Phecode description
Musculoskeletal (7)	**Pain in joint**; **Peripheral enthesopathies and allied syndromes**; **Enthesopathy**; Bursitis; Other and unspecified disc disorder; Other disorders of soft tissues; Juvenile osteochondrosis
Digestive (6)	Hemorrhage of rectum and anus; Symptoms involving digestive system; Irritable Bowel Syndrome; Diseases of lips; Anal and rectal polyp; Dysphagia
Neurological (5)	Acute pain; Organic or persistent insomnia; Other headache syndromes; Other peripheral nerve disorders; Migraine
Injuries & poisonings (3)	Sprains and strains; Contusion; Certain early complications of trauma or procedure
Endocrine/metabolic (2)	Abnormal weight gain; Hypercholesterolemia
Mental disorders (1)	Adjustment reaction

Significant associations with excessive scarring ($p < 0.05/1518 = 3.3\times10\text{-}05$), grouped by category (number of phecodes per category shown in brackets). The 20 most significantly associated phecodes are in bold.

Table 4.9: Detailed summary of phecodes significantly associated with excessive scarring status.

Group	Phecode description	OR (95% CI)	P-value	N cases (Excessive scarring)
Circulatory system				
	Angina pectoris	2.00 (1.59, 2.51)	1.38e-09	96
	Nonspecific chest pain	1.48 (1.28, 1.71)	8.21e-08	271
	Other chronic ischemic heart disease, unspecified	1.88 (1.47, 2.40)	2.64e-07	83
	Coronary atherosclerosis	1.87 (1.46, 2.40)	5.51e-07	79
	Heart valve replaced	3.90 (2.25, 6.73)	6.65e-07	14
	Endocarditis	1.95 (1.48, 2.57)	1.47e-06	57
	Orthostatic hypotension	2.42 (1.59, 3.68)	2.40e-05	24
Dermatologic				
	Sebaceous cyst	2.56 (2.17, 3.03)	9.47e-30	181
	Diseases of hair and hair follicles	2.30 (1.94, 2.73)	1.50e-22	169
	Atopic/contact dermatitis due to other or unspecified	1.77 (1.53, 2.04)	1.44e-15	282
	Other local infections of skin and subcutaneous tissue	1.93 (1.63, 2.29)	1.55e-14	168
	Seborrheic keratosis	1.90 (1.60, 2.26)	7.53e-14	175
	Actinic keratosis	2.03 (1.66, 2.47)	9.91e-13	126
	Acne	2.42 (1.89, 3.11)	1.32e-12	73
	Rash and other nonspecific skin eruption	1.75 (1.49, 2.05)	2.81e-12	202
	Other hypertrophic and atrophic conditions of skin	2.00 (1.62, 2.48)	5.14e-11	102
	Hyperhidrosis	2.54 (1.86, 3.45)	1.67e-09	45
	Disturbance of skin sensation	1.80 (1.48, 2.19)	1.78e-09	123
	Carbuncle and furuncle	1.88 (1.52, 2.33)	4.20e-09	100

Statistical significance is declared at $P < 0.05/1518$.

Group	Phecode description	OR (95% CI)	P-value	N cases (Excessive scarring)
	Pruritus and related conditions	1.74 (1.43, 2.11)	1.04e-08	126
	Diseases of nail, NOS	2.00 (1.50, 2.65)	1.19e-06	53
Digestive				
	Hemorrhage of rectum and anus	1.67 (1.38, 2.02)	6.63e-08	129
	Symptoms involving digestive system	1.59 (1.31, 1.93)	1.96e-06	123
	Irritable Bowel Syndrome	1.54 (1.28, 1.85)	2.98e-06	142
	Diseases of lips	2.54 (1.69, 3.83)	4.83e-06	25
	Anal and rectal polyp	1.97 (1.44, 2.70)	1.67e-05	43
	Dysphagia	1.84 (1.38, 2.46)	2.14e-05	52
Endocrine/metabolic				
	Abnormal weight gain	3.97 (2.37, 6.66)	9.32e-08	16
	Hypercholesterolemia	1.46 (1.22, 1.74)	1.79e-05	176
Genitourinary				
	Mastodynia	2.14 (1.74, 2.64)	2.54e-13	114
	Hypertrophy of breast (Gynecomastia)	3.82 (2.46, 5.91)	9.53e-10	22
	Inflammatory disease of breast	2.51 (1.83, 3.45)	7.43e-09	43
	Cystitis	1.78 (1.43, 2.21)	9.90e-08	100
	Frequency of urination and polyuria	1.95 (1.51, 2.52)	1.96e-07	67
	Lump or mass in breast	1.71 (1.39, 2.10)	2.29e-07	115
	Irregular menstrual cycle	1.84 (1.43, 2.36)	1.00e-06	77
	Excessive or frequent menstruation	1.58 (1.30, 1.93)	2.89e-06	143
	Ovarian cyst	1.98 (1.47, 2.66)	3.71e-06	51
	Noninflammatory disorders of vagina	1.67 (1.33, 2.09)	6.51e-06	93

Statistical significance is declared at $P < 0.05/1518$.

Group	Phecode description	OR (95% CI)	P-value	N cases (Excessive scarring)
	Noninflammatory disorders of vulva and perineum	2.08 (1.50, 2.89)	8.35e-06	40
	Premenstrual tension syndromes	2.06 (1.48, 2.86)	1.05e-05	41
	Urinary tract infection	1.42 (1.21, 1.67)	1.52e-05	200
	Renal failure NOS	2.62 (1.66, 4.15)	2.59e-05	20
	Inflammatory diseases of uterus, except cervix	4.14 (2.09, 8.20)	3.13e-05	9
Infectious diseases				
	Postoperative infection	2.61 (2.00, 3.40)	4.05e-13	63
	Tuberculosis	1.55 (1.35, 1.78)	4.70e-10	302
	Viral warts & HPV	1.76 (1.46, 2.14)	3.62e-09	126
	Dermatophytosis of nail	1.80 (1.47, 2.20)	5.68e-09	113
	Sexually transmitted infections (not HIV or hepatitis)	1.70 (1.41, 2.06)	1.76e-08	134
	Dermatophytosis	1.69 (1.40, 2.05)	4.23e-08	128
	Candidiasis	1.57 (1.31, 1.88)	5.70e-07	153
	Herpes simplex	1.68 (1.35, 2.08)	1.57e-06	97
	Viral infection	1.49 (1.25, 1.78)	5.90e-06	156
	Mycoses	1.87 (1.41, 2.47)	9.33e-06	55
	Gram negative septicemia	3.84 (2.02, 7.32)	2.90e-05	10
Injuries & poisonings				
	Sprains and strains	2.72 (1.91, 3.89)	2.01e-08	33
	Contusion	1.93 (1.45, 2.57)	4.70e-06	52
	Certain early complications of trauma or procedure	2.23 (1.52, 3.29)	3.25e-05	28
Mental disorders				
	Adjustment reaction	1.90 (1.51, 2.38)	1.45e-08	88

Statistical significance is declared at $P < 0.05/1518$.

Group	Phecode description	OR (95% CI)	P-value	N cases (Excessive scarring)
Musculoskeletal				
	Pain in joint	1.84 (1.61, 2.10)	1.87e-20	440
	Peripheral enthesopathies and allied syndromes	1.75 (1.50, 2.03)	6.53e-14	247
	Enthesopathy	1.72 (1.48, 2.01)	2.28e-12	221
	Bursitis	1.89 (1.50, 2.38)	4.40e-08	84
	Other and unspecified disc disorder	1.44 (1.23, 1.67)	1.65e-06	236
	Other disorders of soft tissues	1.63 (1.31, 2.03)	9.94e-06	93
	Juvenile osteochondrosis	3.62 (2.01, 6.51)	1.24e-05	12
Neoplasms				
	Other non-epithelial cancer of skin	2.89 (2.36, 3.54)	2.03e-25	120
	Melanomas of skin	3.17 (2.24, 4.50)	3.33e-11	35
	Malignant neoplasm, other	2.05 (1.61, 2.61)	2.78e-09	77
	Lipoma	1.95 (1.50, 2.54)	3.28e-07	63
	Other benign neoplasm of connective and other soft tissue	3.30 (1.91, 5.68)	1.14e-05	14
	Benign neoplasm of lip, oral cavity, and pharynx	3.61 (2.01, 6.49)	1.21e-05	12
	Malignant neoplasm of female breast	1.75 (1.35, 2.28)	2.04e-05	66
	Nevus, non-neoplastic	2.59 (1.64, 4.08)	2.95e-05	20
Neurological				
	Acute pain	1.87 (1.53, 2.28)	4.02e-10	117
	Organic or persistent insomnia	2.08 (1.63, 2.66)	2.44e-09	73
	Other headache syndromes	1.59 (1.34, 1.88)	4.45e-08	176
	Other peripheral nerve disorders	1.60 (1.31, 1.94)	1.80e-06	124

Statistical significance is declared at $P < 0.05/1518$.

Group	Phecode description	OR (95% CI)	P-value	N cases (Excessive scarring)
	Migraine	1.53 (1.27, 1.84)	6.29e-06	137
Respiratory				
	Acute upper respiratory infections of multiple or unspecified sites	1.66 (1.45, 1.89)	2.53e-14	402
	Cough	1.63 (1.42, 1.88)	1.53e-12	321
	Acute sinusitis	1.79 (1.52, 2.12)	2.73e-12	181
	Other diseases of respiratory system, NEC	1.57 (1.37, 1.80)	1.67e-11	380
	Acute pharyngitis	1.66 (1.43, 1.94)	2.51e-11	232
	Acute laryngitis and tracheitis	1.89 (1.52, 2.35)	6.22e-09	96
	Influenza	1.76 (1.41, 2.20)	2.57e-07	93
	Shortness of breath	1.69 (1.36, 2.10)	1.23e-06	100
	Pneumonia	1.63 (1.31, 2.03)	9.67e-06	93
	Pneumonia due to fungus (mycoses)	1.89 (1.41, 2.53)	1.30e-05	51
	Chronic sinusitis	1.53 (1.26, 1.87)	1.57e-05	118
	Chronic pharyngitis and nasopharyngitis	1.67 (1.31, 2.15)	3.26e-05	71
Sense organs				
	Conjunctivitis, infectious	1.84 (1.57, 2.16)	1.66e-14	205
	Dizziness and giddiness (Light-headedness and vertigo)	1.70 (1.44, 2.01)	1.29e-10	187
	Inflammation of eyelids	1.78 (1.48, 2.15)	4.98e-10	137
	Eustachian tube disorders	1.97 (1.57, 2.47)	2.93e-09	86
	Disorders of lacrimal system	1.84 (1.45, 2.33)	2.73e-07	82
	Otalgia	1.69 (1.37, 2.08)	4.84e-07	106
	Otitis externa	1.51 (1.27, 1.79)	1.19e-06	173
	Infection of the eye	1.91 (1.46, 2.49)	1.21e-06	61

Statistical significance is declared at $P < 0.05/1518$.

Group	Phecode description	OR (95% CI)	P-value	N cases (Excessive scarring)
	Hearing loss	1.62 (1.33, 1.97)	1.38e-06	119
	Conjunctivitis, noninfectious	2.46 (1.65, 3.68)	7.13e-06	26
	Tinnitus	1.74 (1.35, 2.25)	1.37e-05	67
Symptoms				
	Malaise and fatigue	1.90 (1.61, 2.24)	9.42e-15	187
	Back pain	1.66 (1.46, 1.90)	1.45e-14	397
	Cervicalgia	1.82 (1.55, 2.13)	5.58e-14	205
	Abdominal pain	1.55 (1.36, 1.77)	3.33e-11	412
	Nausea and vomiting	1.69 (1.37, 2.07)	4.20e-07	108
	Swelling of limb	1.95 (1.43, 2.65)	1.38e-05	46
	Sciatica	1.48 (1.23, 1.78)	2.09e-05	141
	Edema	1.77 (1.35, 2.33)	2.63e-05	61

Statistical significance is declared at $P < 0.05/1518$.

Other previously-reported associations that did not meet the selection criteria for specific analysis were explored, including obesity, osteoporosis, skin cancers, pancreatic cancer, migraine and asthma. Of these, (phenome-wide) statistically-significant associations were observed for skin cancers (melanoma, OR 3.17, $p=3.33 \times 10^{-11}$) and migraine (OR 1.53, $p=6.29 \times 10^{-6}$; Table 4.10).

Table 4.10: Phecodes of previously reported disease associations with excessive scarring status.

Group	Phecode description	OR (95% CI)	P-value	N cases (Excessive scarring)
Circulatory system				
	Essential hypertension	1.26 (1.08, 1.46)	2.05e-03	357
Dermatologic				
	Atopic/contact dermatitis due to other or unspecified	1.77 (1.53, 2.04)	1.44e-15	282
Endocrine/metabolic				
	Vitamin D deficiency	1.47 (1.08, 2.00)	1.34e-02	50
	Obesity	1.27 (1.03, 1.58)	2.34e-02	139
Musculoskeletal				
	Osteoporosis NOS	1.17 (0.85, 1.60)	3.33e-01	45
Neoplasms				
	Other non-epithelial cancer of skin	2.89 (2.36, 3.54)	2.03e-25	120
	Melanomas of skin	3.17 (2.24, 4.50)	3.33e-11	35
	Uterine leiomyoma	1.18 (0.93, 1.49)	1.68e-01	87
	Pancreatic cancer	0.58 (0.08, 4.32)	5.89e-01	1
Neurological				
	Migraine	1.53 (1.27, 1.84)	6.29e-06	137
Respiratory				
	Asthma	1.17 (0.99, 1.39)	6.32e-02	169

Statistical significance is declared at $P < 0.05/1518$

4.4 Discussion

To date there have been few large-scale association studies for excessive scarring(Kwon et al., 2021; Y.-Y. Lu et al., 2021; Lu et al., 2018; Sun et al., 2014; Yu et al.,

2013). This study aimed to both validate previously-studied associations (hypertension(Adotama et al., 2015; Lawton Snyder et al., 1996; Rutherford and Glass, 2017), uterine leiomyoma(Harmon et al., 2013; Sun et al., 2014), vitamin D deficiency(El Hadidi et al., 2021; Yu et al., 2013), and atopic eczema(Kwon et al., 2021; Lu et al., 2018)) as well as scan for excessive scarring associations across the phenome. The ethnicity-specific analysis represents the first comprehensive study in white people with excessive scarring, the most-represented ethnic group in UKB.

In the trans-ethnic UKB population, associations with three (hypertension, vitamin D deficiency and atopic eczema) of the four primary comorbidities previously studied two or more times were investigated. Only the associations with hypertension and atopic eczema were statistically significant. However, the association with hypertension was attenuated after adjusting for additional risk factors (BMI, Townsend Deprivation Index, smoking-status, diabetes or hyperlipidaemia). This suggests mediation by indirect effects resulting from differences in these risk factors between people with and without excessive scarring.

Subgroup analyses revealed ethnic variation in disease risk; the associations with hypertension and uterine leiomyoma were unique to black participants (hypertension, OR 2.05, p=0.019; leiomyoma, OR 1.93, p=0.05) whereas the association with vitamin D deficiency was unique to Asian participants (OR 2.24, p=0.006). The disproportionate burden of these diseases within the respective ethnic groups(Dustan, 1995; Pavone et al., 2018; Stewart et al., 2017; Williams et al., 2016) and the relatively small sample sizes in this study make a robust interpretation challenging; however it is possible that there are ethnicity-specific risk determinants shared by these pathologies. For example, vascular dysfunction is thought to contribute to the severe hypertension and hypertensive heart failure specifically affecting the black population(Nayak et al., 2020; Sinha et al., 2021). Abnormal endothelial function and microvascular architecture are also observed in excessive scarring (in particular, keloids) and uterine leiomyoma(Kurokawa et al., 2010; Matsumoto et al., 2020; Noishiki et al., 2019; Tal and Segars, 2013). Replication of these findings in a larger

cohort may make a case for the early identification of cardiovascular disease in black individuals with excessive scarring and/or uterine leiomyoma.

The only association that showed nominally significant evidence for association in multiple ethnic groups was atopic eczema, previously reported in Taiwanese(Lu et al., 2018) and Korean(Kwon et al., 2021) populations. Atopic eczema skin is more likely to be excoriated and scars may thus be a commoner phenomenon. However, this finding also adds epidemiological support to the hypothesis that the Th2 inflammatory axis contributes to excessive scarring pathogenesis(Wu et al., 2020). A cross-talk between the Th2 pathway and fibrosis was suggested through the discovery of interleukin (IL) 4 and IL13 acting on the matricellular protein, periostin, to increase TGFβ1 signalling(Maeda et al., 2019) and increasing collagen production(Nguyen et al., 2019; Oriente et al., 2000). Molecular profiling of keloid skin highlighted the Th2 axis as part of a broad immune dysregulation signature(Wu et al., 2020). Case reports of keloid shrinkage(Diaz et al., 2020) and symptom improvement(Wong and Song, 2021) following treatment with the recombinant anti-IL4/IL13 antibody, dupilumab, have resulted in a clinical trial investigating its efficacy as treatment for keloid (NCT04988022).

Previous non-genetic application of PheWASs(Boland et al., 2013; Liao et al., 2013; Warner et al., 2013; Zhang et al., 2020) have been based only on ICD diagnostic codes. This would have excluded a large proportion of excessive scarring cases only identified through primary care codes. Through the comprehensive strategy to maximize identification of people with excessive scarring, numerous significant disease associations were highlighted, potentially indicating an increased risk of poorer health outcomes.

The frequent female genitourinary disease associations may be explained by the over-representation of female participants within the excessive scar-affected group. Nonetheless, adjustments were made for sex and it has been suggested that sex hormones may play a role in excessive scarring pathophysiology(Ibrahim et al., 2020; Moustafa et al., 1975; Noishiki et al., 2019). The highly-significant associations with dermatological conditions and neoplasms may represent true predispositions or

more likely, reflect ascertainment bias (i.e. if a patient presents with a dermatological condition or is reviewed post-surgically, it is more likely that a scar-related diagnosis will be recorded).

Of the dermatological associations, diseases of the pilosebaceous unit (sebaceous cyst, diseases of hair/hair follicles, acne) support the 'sebum hypothesis'(Fong and Bay, 2002) which is based on high sebaceous gland density observed particularly in keloid-prone skin(Limandjaja et al., 2020b) and sebum being intrinsically pro-inflammatory(Yagi et al., 1979). The association of keloid with skin cancer has been previously reported(C.-C. Lu et al., 2021). Although plausible reasons have been put forward including similar bioenergetics (reliance on glycolysis)(Onoufriadis et al., 2018; Vincent et al., 2008) and signalling pathways including TGFβ/*Smad*(He et al., 2009; Liu et al., 2016; Zhao et al., 2018) and Wnt/β-Catenin(Chaudet et al., 2020; Driskell and Watt, 2015; Igota et al., 2013; Koni et al., 2020), this finding is interpreted cautiously, again considering the risk of ascertainment bias.

The associations with musculoskeletal disorders (enthesopathy, pain in joint, back pain, cervicalgia) may support the observation of chondrogeneic misdifferentiation in keloids(Barallobre-Barreiro et al., 2019) and the shared significance of TGFβ in joint pathologies(Clayton et al., 2020; Wang and Zhang, 2018). Interestingly, associations with pain symptoms spanned disease categories (non-specific chest pain, irritable bowel syndrome, mastodynia, acute pain, headache syndromes, pain in joint, back pain, cervicalgia). Pain is known to debilitate some keloids sufferers(Hawash et al., 2021); whether there is shared underlying biopsychosocial dysfunction with other pain entities or whether they may be mutually reinforcing is speculative. Nonetheless, chronic pain represents a major global burden of disease(Perrot et al., 2019) and proactive identification of these conditions may aid patient counselling and treatment decisions.

Finally, whether an individual whose skin scars excessively is at risk of excessive internal scarring remains unanswered. In our study, the association of peritoneal

adhesions with excessive scarring carried an OR of 3.68 (p=0.0001), which is intriguing but this is based only on 617 cases of peritoneal adhesions, nine of whom had excessive scarring.

4.5 Limitations

Although this study utilized a large biobank cohort, a relatively small sample size of excessive scarring cases were identified (n=972). Small sample sizes are also subject to confounding, which in this study, was controlled for by regression adjustment. An alternative statistical method to reduce confounding is matching (e.g. propensity score matching(Austin, 2011)) where cases and controls share potential confounding characteristics. However, the matching procedure may reduce statistical power of studies and have been shown to yield identical results as if matching was not performed.(Faresjö and Faresjö, 2010) This is relevant when attempting to dissect differences between ethnic groups, as there is less statistical power to detect significant associations in groups with lower sample numbers.

As the majority of participants report white ethnicity, the main findings, particularly from the PheWAS, may not be generalizable to other ethnic groups for whom excessive scarring is a more prominent issue. Sample size was also a limitation for the detection of associations for some phecode categories, as was the case for peritoneal adhesions.

From the PheWAS, there are a few further discussion points. As a high throughput phenotyping method was used, potential code conflicts may have arisen. For example, if an individual has diagnostic codes for both type 1 and type 2 diabetes in their health records, they would have been labelled as a 'case' for both diagnoses. Further investigation for specific disease associations of interest should therefore employ validated phenotyping algorithms(Bastarache, 2021). The lack of significant associations with other fibrotic/scarring comorbidities (e.g. lung or liver fibrosis) supports previous reports(Bayat et al., 2005a; Sun et al., 2014) although low case numbers may mean our investigations were insufficiently powered to detect them.

Secondly, it was striking that all significant associations were positive (i.e. increased prevalence of comorbidities in people with excessive scarring), potentially a result of coverage bias whereby participants with more complete coverage of linked health data may be more likely to have a record of excessive scar diagnosis as well as a diagnosis of any other comorbidity. This may mean that the effect sizes are overestimated; nonetheless the relative order of the associations remains informative.

5

Determining genetic associations of excessive scarring

Contents

This chapter presents results from a GWAS for excessive scarring in the UKB cohort and a meta-analysis of the UKB study with a GWAS of excessive scarring in the Finnish founder population (FinnGen).

GWAS is a widely used analysis that aims to identify genotype associations with phenotypes/diseases/traits by testing for differences in the variant allele frequency between ancestrally-similar individuals. Over the last two decades, GWASs have yielded numerous genetic associations for many heritable phenotypes/traits; the

biological insight gained has applications ranging from disease risk prediction to therapeutic development (e.g. implication of IL12/23 in Crohn's disease(Wang et al., 2009), supporting clinical trials for drugs targeting this pathway(Moschen et al., 2018)). The standard approach tests for association between single variants and the trait; typically, blocks of correlated variants (genomic risk loci) may exhibit association with the trait, as a result of linkage disequilibrium, LD, the non-random association between variants in close proximity due to lower recombination rates(Hill and Robertson, 1968). By being in LD with the true causal variant, numerous non-causal variants can be associated with the trait studied. Various bioinformatic methods exist to discover the true causal variant(s) and prioritise likely affected gene(s) to derive relevant biological conclusions. An alternative approach attributes variation to genes and evaluates their cummulative association with the trait. This approach increases study power (due to lower penalty for multiple testing) and adds a layer of functional interpretation, but is limited in the analysis of the non-coding genome(Ma et al., 2021). Although mainly focused on single variant association analysis, the latter approach is also explored in this chapter; findings and limitations will be discussed.

5.1 Introduction

Even though excessive scarring (i.e. keloid and/or hypertrophic scars) can afflict all ethnicities, the focus of previous genetic studies has been on Asian and black individuals, in whom the prevalence is greater. Three GWASs and one cross-population meta-analysis* have been published to date, reporting excessive scarring susceptibility(Ishigaki et al., 2020; Nakashima et al., 2010; Sakaue et al., 2021; Zhu et al., 2013); two shared risk loci were identified: 1q41 and 15q21.3. The Japanese studies identified additional loci in 3q22.3-23(Ishigaki et al., 2020; Nakashima et al., 2010;

*The meta-analysis was a systematic analysis of over 200 phenotypes and combined data from Biobank Japan, UKB and FinnGen. The clinical code used to define cases for this study (ICD10 code L91.0) only identified 214 keloid cases from UKB and 454 cases from FinnGen.

Sakaue et al., 2021) and 1q32.1(Ishigaki et al., 2020) whereas the meta-analysis identified 2q37.3 as a novel risk locus (although no genome-wide significant association was identified in any of the individual datasets analysed, and the signal appears to be largely driven by Japanese cases)(Sakaue et al., 2021). Admixture mapping analysis in a black cohort (122 cases and 356 controls) also identified a variant within 15q21.2-22.3(Velez Edwards et al., 2014), suggesting that the chromosome 15 susceptibility locus may not be ethnicity specific.

In this chapter, a GWAS was performed in white British individuals with a diagnosis of keloids and hypertrophic scars (i.e. excessive scarring) within the UKB cohort. This was repeated in a sensitvity analysis using a strict definition of keloid. Whereas previously published GWASs (including the recent meta-analysis(Sakaue et al., 2021)) have identified excessive scarring cases using secondary care records alone, both primary and secondary care records were utilised for the UKB analysis in order to maximise statistical power for discovery. To systematically analyse effect sizes with increased power, a meta-analysis was performed, combining the UKB study with results from FinnGen, a Finnish genome and national health register study of 1932 clinical endpoints, including keloids and hypertrophic scars. The genotyping, imputation and primary quality control of the UKB data were performed by UKB. I conducted the cohort selection, subsequent quality control, GWAS (UKB), and downstream functional analyses with the exception of the the Bayesian fine-mapping of putative variants and the meta-analysis which were performed by my collaborators Dr Jake Saklatvala and Dr Nick Dand.

5.2 Cohort summary

As detailed in Chapter 2, the UKB is a prospective cohort of over 500000 individuals recruited between the ages of 40 and 69 from 2006 and 2010 across the UK, in whom rich phenotyping and genotyping data were collected. Genotyping was performed using the Affymetrix UK BiLEVE Axiom Array (807411 markers on 49950 individuals) and the Affymetrix UK Biobank Axiom Array (825925 markers on 438427

individuals)(Bycroft et al., 2018). Both platforms shared 95% of common genetic markers, and quality control and imputation (the process of predicting genotypes that are not directly assayed to boost the number of SNPs that can be tested for association) were performed jointly(Bycroft et al., 2018). Specifically, genotype imputation was based on the Haplotype Reference Consortium, UK10K and 1KGP phase 3 reference panels, resulting in 93095623 genetic markers in 487442 individuals.

Analyses were restricted to individuals with relatively homogeneous ancestry, excluding those with poor sample quality. Of the 502411 participants available for study, 402628 remained after excluding individuals of non-white-British ancestry, with sex-discordance, with excess genotype heterozygosity or missingness and with 'excess relatedness' as defined by the UKB central team and provided as QC metrics. A further 213973 individuals without linked primary care records were removed to reduce misclassification bias (e.g. individuals with a diagnosis of keloids only in a primary care setting being misclassified as controls).

Of the remaining participants, a total of 843 participants were found to have a record of excessive scarring based on secondary (ICD versions 9 and 10) and primary (Read versions 2 and 3) care diagnosis codes, compared to 187812 without; amongst the cases, 509 had a diagnosis code strictly for keloid and 334 had a diagnosis code of keloid and/or hypertrophic scar).

Table 5.1 shows the baseline characteristics for the excessive scarring and comparator groups. In the excessive scar-affected group, there was a higher proportion of females than in the unaffected group (66% versus 54%, $p<0.001$) and a lower mean body mass index (28.0 versus 27.5, $p=0.014$).

Table 5.1: Baseline characteristics for the excessive scarring and comparator groups.

Characteristic	**Case, N = 843[1]**	**Control, N = 187,812[1]**	**p-value[2]**
Age	64 (8)	65 (8)	<0.001
Sex			<0.001
Female	556 (66%)	102,111 (54%)	
Male	287 (34%)	85,701 (46%)	
Townsend Deprivation Index	-1.59 (3.00)	-1.54 (2.92)	0.6
Ever smoked	481 (57%)	111,941 (60%)	0.2
Body mass index	28.0 (5.2)	27.5 (4.8)	0.014

[1]Mean (SD); n (%)

[2]Welch Two Sample t-test; Pearson's Chi-squared test

5.3 Genome-wide association analysis

GWAS analyses were performed for two keloid definitions: excessive scarring (keloid and hypertrophic scar diagnosis codes) and keloid (keloid diagnosis codes only), using sex, genotype array and the first five ancestry principal components as covariates, in a logistic mixed-effects model (which accounts for relatedness and population structure). Genomic inflation factor calculation and LD score regression coefficients suggested non-significant genomic inflation due to population stratification (Table 5.2). The corresponding Quantile-quantile (QQ)-plots are shown in Figures 5.1 and 5.2.

Table 5.2: Genomic inflation analyses.

Analysis	Cases	Controls	Genomic inflation factor	LDSC intercept (SD)
Excessive scarring	843	187,812	1.016	0.989 (0.007)
Keloid	509	187,812	1.005	0.989 (0.007)

LDSC, linkage disequilibrium score regression; SD, standard deviation

Using LDSR to estimate the heritability explained by all the SNPs tested (the amount of variance on the risk of excessive scarring explained by this GWAS), a heritability index of 0.0068(0.0023) was identified for excessive scarring and 0.0047(0.0024) for keloid. Based on the low sample sizes, it is likely that the study was insufficiently powered to detect significant heritability estimates. As a result, these low heritability indices precluded subsequent estimation of genetic correlation with other traits.

The excessive scarring analysis identified two genome-wide significant loci: 1q41 and 8q21.3 (Figure 5.1), comprising 186 SNPs in LD with two lead SNPs (rs10863683, OR=1.48 (1.34-1.63), p=1.32x10^{-15}; rs574775574, OR=0.57 (0.47-0.7), p=1.25x10^{-8}) associated with excessive scarring. However, as the 8q21.3 locus comprised only a single SNP, this was cautiously interpreted. Notably, the 1q41 locus is consistent with previous GWASs, albeit with a different lead SNP(Nakashima et al., 2010;

Shirai et al., 2022; Zhu et al., 2013). The regional plot for the 1q41 locus is shown in Figure 5.3. This association was confirmed in the strict keloid analysis (Figure 5.2).

5.4 Genome-wide assoiation meta-analysis

A meta-analysis was performed incorporating data from FinnGen (data release 6, 2021), which comprises an independent European ancestry cohort, made up of a relatively young, isolated population with a unique genetic makeup. The FinnGen GWAS was based on genotypes imputed using a Finnish population-specific reference panel and restricted to unrelated individuals with Finnish ancestry after removing population outliers. A broad 'hypertrophic scar' diagnosis was used to define cases, using ICD (version 8, 9 and 10) codes which included keloids - 1126 cases and 291889 controls were included for association analysis. The SAIGE mixed model logistic regression method was used to test for associations, with sex, age, genotyping batch and 10 principal components as covariates.

With the resulting 1969 cases and 479701 controls, significant associations were identified in three independent loci, previously only established in Asian populations: 1q32.1/chr1:201437832-201458214 (most significant SNP, rs35383942, OR 1.46, p=6.991x10^{-11}) and 1q41/chr1:222243525-222427283 (rs10863683, OR 1.45, p=3.034x10^{-25}) at genome-wide significance and 15q21.3/chr15:56089229-56186767 (rs60890210, OR 1.24, p=6.723x10^{-08}) approaching genome-wide significance (Table 5.3, Figure 5.4). Figures 5.5 shows the regional plots for the three loci. The chromosome 3 risk locus previously reported in the Japanese population was not replicated (published variant, rs646315, p=0.4244; not significant in either UKB or FinnGen). A comparison of the meta-analysis results with previously published risk variants is summarised in Table 5.4.

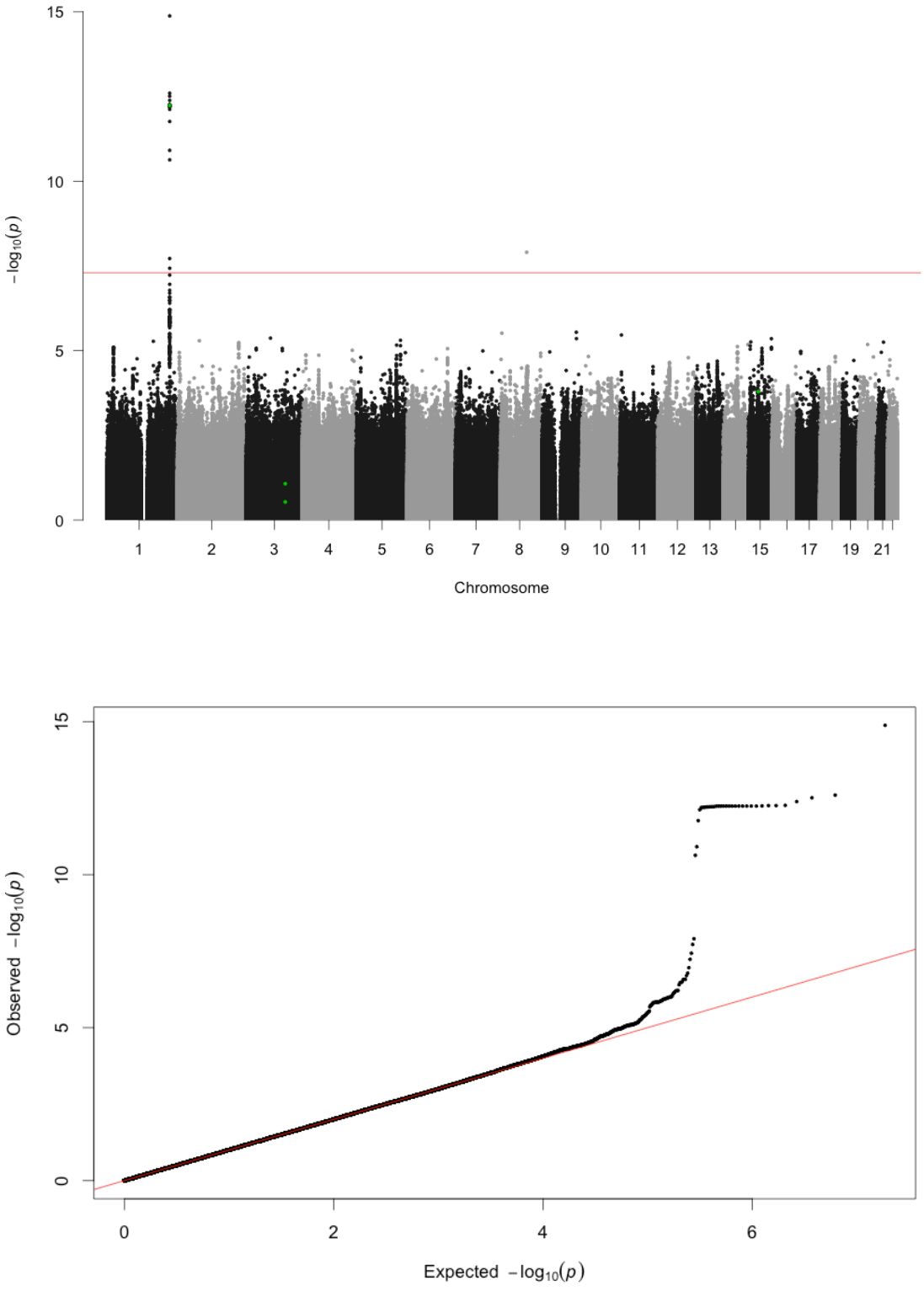

Figure 5.1: Top panel. Manhattan plot showing genetic associations for excessive scarring in the UKB. Red line: Bonferroni-corrected significance threshold. Green dots: Previously-reported keloid risk variants. Bottom panel. QQ-plot.

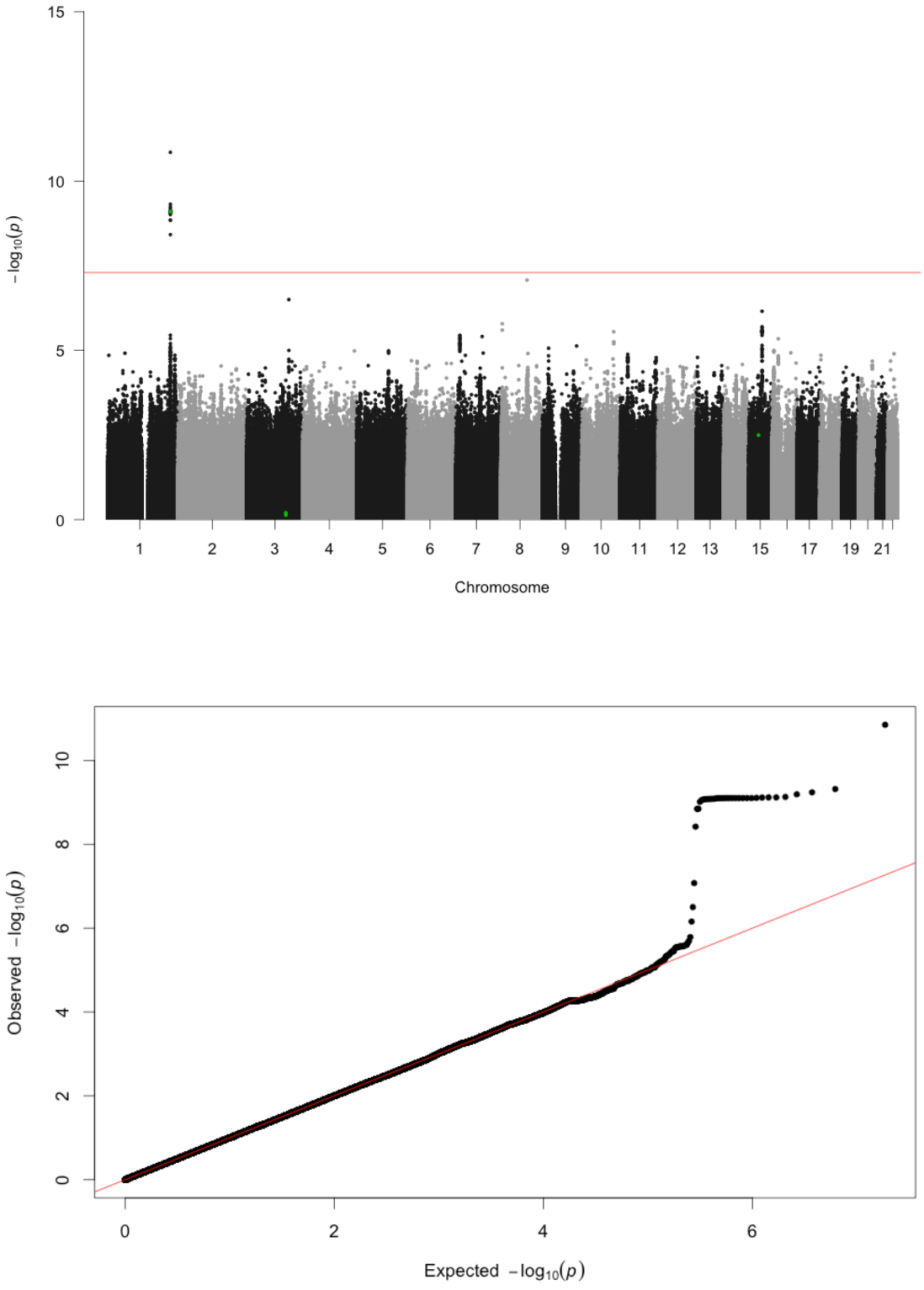

Figure 5.2: Top panel. Manhattan plot showing genetic associations for keloids in the UKB. Red line: Bonferroni-corrected significance threshold. Green dots: Previously-reported keloid risk variants. Bottom panel. QQ-plot.

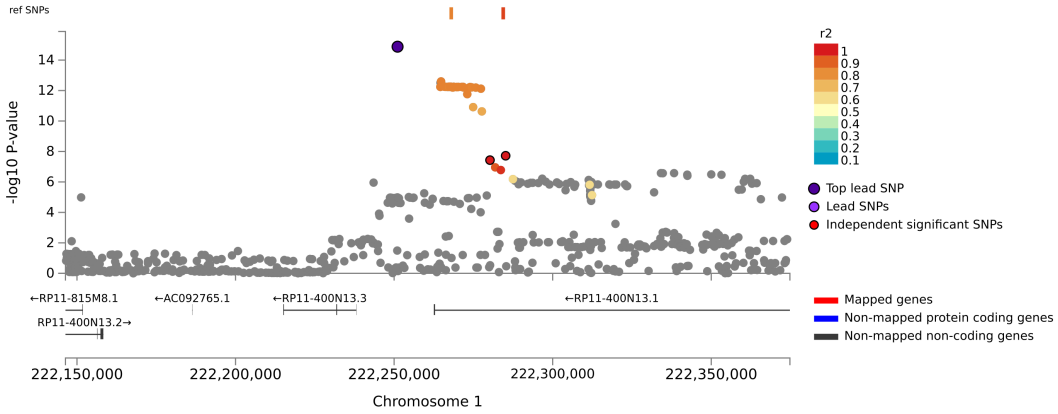

Figure 5.3: Regional plot for the 1q41 locus with the excessive scarring GWAS p-value generated within FUMA (showing the range of LD across the region). Independent significant SNPs (defined by $P<5x10^{-8}$ and $r^2<0.6$) with $r^2 > 0.1$ are assigned to the same risk locus. SNPs are colour-coded based on the highest r^2 (a marker of LD). The independent significant SNPs and the top lead SNP which has the minimum p-value in the locus are indicated. SNPs with $r^2<0.1$ are coloured in grey.

Table 5.3: Independent risk loci and lead SNPs from meta-analysis.

				UK Biobank		FinnGen		Meta-analysis	
SNP	**Locus**	**Ref**	**Alt**	**OR**	**P**	**OR**	**P**	**OR**	**P**
rs10863683	1q41	C	G	0.68 (0.62; 0.75)	4.60e-15	0.7 (0.64; 0.78)	8.88e-12	0.69 (0.64; 0.74)	3.03e-25
rs35383942	1q32.1	C	T	1.35 (1.13; 1.63)	1.04e-03	1.53 (1.32; 1.78)	9.83e-09	1.46 (1.3; 1.64)	6.99e-11
rs60890210	15q21.3	G	A	1.21 (1.08; 1.34)	4.60e-04	1.28 (1.14; 1.44)	3.06e-05	1.24 (1.14; 1.34)	6.72e-08

SNP, single nucleotide polymorphism; Ref, reference (non-effect) allele; Alt, alternate (effect) allele; OR, odds ratio (95% confidence interval)

Table 5.4: Comparison of meta-analysis results with previously published keloid lead variants.

				Published GWAS		UK Biobank		FinnGen		Meta-analysis	
SNP	**Locus**	**Ref**	**Alt**	**OR**	**P**	**OR**	**P**	**OR**	**P**	**OR**	**P**
rs192314256[a]	1q32.1	T	C	9.56	3.3e-20	—	—	—	—	—	—
rs2378519[a]	1q41	G	A	0.54	6.7e-27	0.7 (0.63; 0.77)	1.1e-12	0.73 (0.66; 0.81)	5.4e-09	0.71 (0.66; 0.77)	4.4e-20
rs11293015[d]	1q41	GT	G	0.6	3.7e-26	—	—	—	—	—	—
rs873549[bc]	1q41	C	T	0.56	5.9e-23	0.7 (0.63; 0.77)	1.1e-12	0.73 (0.66; 0.82)	6.0e-09	0.71 (0.66; 0.77)	4.7e-20
rs1442440[c]	1q41	T	C	0.56	9.9e-18	0.88 (0.79; 0.97)	9.7e-03	0.92 (0.83; 1.01)	7.4e-02	0.9 (0.84; 0.96)	2.1e-03
rs74983791[d]	2q37.3	T	C	1.31	3.7e-08	1.04 (0.74; 1.48)	8e-01	—	—	—	—
rs1511412[b]	3q23	A	G	0.53	2.3e-13	0.93 (0.8; 1.09)	3.8e-01	0.97 (0.8; 1.16)	7.2e-01	0.95 (0.84; 1.07)	3.6e-01
rs940187[b]	3q23	T	C	0.51	1.8e-13	0.9 (0.8; 1.02)	8.5e-02	0.91 (0.8; 1.04)	1.5e-01	0.9 (0.83; 0.99)	2.4e-02
rs646315[a]	3q23	G	T	2.05	4.2e-12	1.06 (0.92; 1.23)	4.2e-01	1.03 (0.85; 1.23)	7.8e-01	1.05 (0.93; 1.17)	4.2e-01
rs8032158[b]	15q21.3	T	C	1.51	6.0e-13	1.2 (1.09; 1.33)	2.6e-04	1.21 (1.09; 1.34)	1.5e-04	1.21 (1.12; 1.3)	1.4e-07
rs2271289[c]	15q21.3	C	T	0.66	1.0e-11	0.9 (0.81; 1)	4.3e-02	0.95 (0.86; 1.05)	3.3e-01	0.93 (0.86; 1)	3.6e-02
rs16976600[a]	15q21.3	C	T	1.43	1.3e-11	1.2 (1.08; 1.32)	4.4e-04	1.2 (1.09; 1.33)	2.1e-04	1.2 (1.12; 1.29)	3.3e-07

SNP, single nucleotide polymorphism; Ref, reference (non-effect) allele; Alt, alternate (effect) allele; OR, odds ratio; [a]PMID 32514122; [b]PMID 20711176; [c]PMID 23667473; [d]PMID 34594039

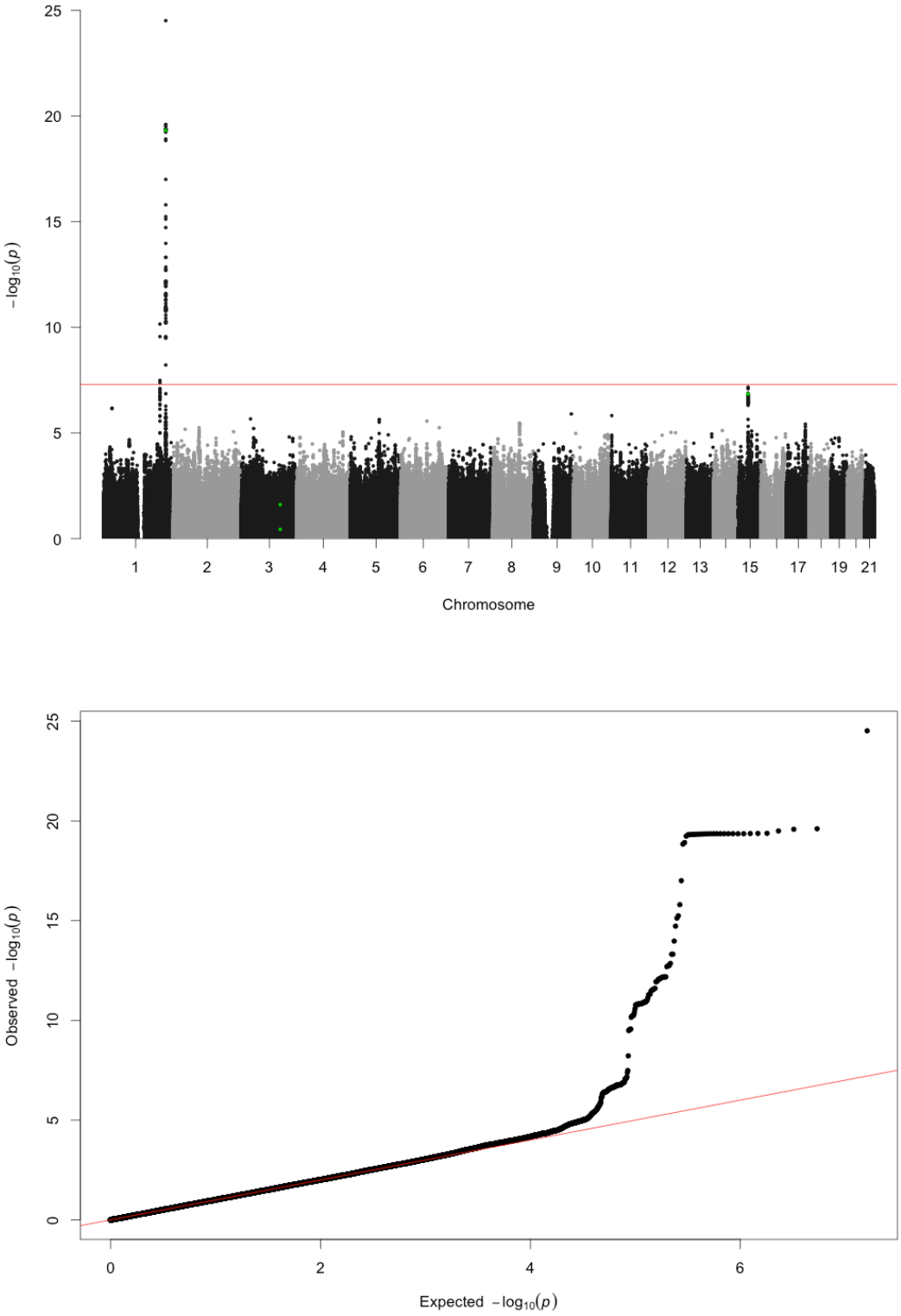

Figure 5.4: Top panel. Manhattan plot showing genetic associations for excessive scarring from the meta-analysis. Red line: Bonferroni-corrected significance threshold. Green dots: Previously-reported keloid risk variants. Bottom panel. QQ-plot.

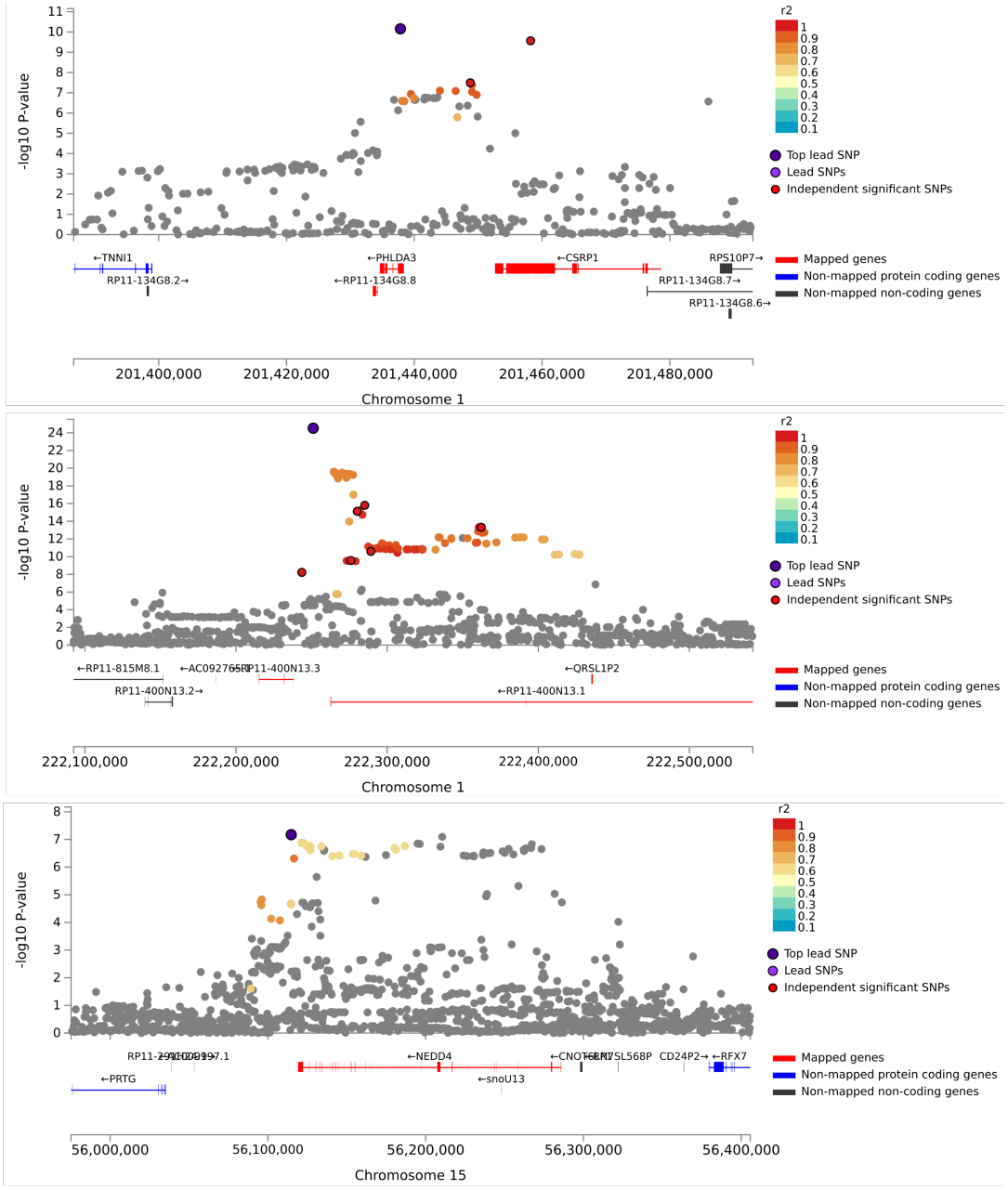

Figure 5.5: Regional plots for the 1q32.1 (top panel), 1q41 (middle panel) and 15q21.3 (bottom panel) loci , with the meta-analysis p-value, generated within FUMA (showing the range of LD across the region). Independent significant SNPs (defined by $P<5x10^{-8}$ and $r^2<0.6$) with $r^2 > 0.1$ are assigned to the same risk locus. SNPs are colour-coded based on the highest r^2 (a marker of LD). The independent significant SNPs and the top lead SNP which has the minimum p-value in the locus are indicated. SNPs with $r^2<0.1$ are coloured in grey.

5.4.1 Credible set fine-mapping

As many SNPs may be associated with a trait by being in LD with the causal SNP and the most significant SNP is not necessarily causal, statistical fine-mapping was performed to identify likely causal SNPs. Bayesian fine-mapping was performed at the three risk loci for UKB, FinnGen and meta-analysis results. The number of variants in the 95% credible set (the smallest set of variants with sum of posterior probabilities of being causal variants $\geq 95\%$) for both datasets and the meta-analysis are summarised in Table 5.5 with the top shared variants and their individual posterior probabilities. For two of the three risk loci, a single variant was identified as the putative causal variant with a posterior probability (PP) > 0.78: rs10863683 (chr1:222251089, PP 0.786) and rs35383942 (chr1:201437832, PP 0.999). Already identified as the lead SNP from the UKB GWAS, rs10863683 is an intergenic SNP with the long non-coding RNA *RP11-400N13.1* as its nearest gene, whereas rs35383942 is a missense variant within *PHLDA3*, which encodes a p53-regulated repressor of AKT(Kawase et al., 2009). Evidence for causality of the lead SNP at 15q21.3 was inconclusive.

Table 5.5: Fine-mapping summary.

		UK Biobank		**FinnGen**		**Meta-analysis**	
Locus	**Top shared BCS variant**	**BCS size**	**PP**	**BCS size**	**PP**	**BCS size**	**PP**
1q32.1	rs35383942	694	0.0036†	20	0.1596†	2	0.7860
1q41	rs10863683	20	0.8408	15	0.8782	1	0.9997
15q21.3	rs60890210	595	0.0049†	1,133	0.0321†	46	0.0625

BCS, Bayesian 95% credible set; PP, posterior probability.

†Bayesian fine-mapping model allocated a higher posterior probability to the scenario where there is no causal association than to any individual variant being causal.

5.4.2 Functional annotation

Using tools embedded within FUMA(Watanabe et al., 2017), an online platform for functional mapping and annotation of genetic variants, a total of 170 candidate causal variants from the meta-analysis (including all SNPs in LD (r^2>0.6) of the identified lead SNPs) were annotated. The annotations included functional consequences of the variants (ANNOVAR), the predicted deleteriousness of the variants (Combined Annotation Dependent Depletion, CADD score), the functionality of the variants in transcriptional regulation (RegulomeDB) and the predicted chromatin states for the variants' genomic segment (ChromHMM). A summary of these annotations for the lead variants is in Table 5.6. Notably, a missense variant within *PHLDA3*, rs35383942, also the variant with a posterior probability of being causal of 0.99, was predicted to have a highly deleterious effect (CADD score 24.3).

Table 5.6: FUMA annotations of lead SNPs.

	rs35383942	**rs10863683**	**rs60890210**
Locus	1q32.1	1q41	15q21.3
MAF	0.05964	0.3111	0.2058
OR (95%CI)	1.46 (1.3 - 1.64)	1.44 (1.35 - 1.55)	1.24 (1.14 - 1.34)
P	6.99e-11	3.03e-25	6.72e-08
Nearest gene	*PHLDA3*	*RP11-400N13.1*	*NEDD4*
ANNOVAR annotation	exonic	intergenic	intergenic
CADD	24.3	0.158	0.122
RDB	2a	7	6
ChrState (min)	1	5	5
ChrState (common)	1	15	15

MAF, minor allele frequency; OR (95% CI), odds ratio (95% confidence interval); CADD, Combined Annotation Dependent Depletion score; RDB, RegulomeDB score (1a to 7, 1a being the score for SNPs with the most biological evidence to be a regulatory element); ChrState, 15-core chromatin state across 127 tissue/cell types predicted by ChromHMM, consisting of eight active states (1-8) associated with expressed genes and seven repressed states (9-15).

To identify additional putative causal genes at the risk loci, a combination of bioinformatics approaches were taken: gene-mapping of independent SNPs based on physical proximity (positional mapping), and eQTLs in skin, dermal fibroblasts and blood (eQTL mapping). Genes mapped by eQTL were analysed using a Bayesian test for colocalisation, to evaluate the probability of evidence of a single shared causal variant between excessive scarring risk and transcript expression.

Positional and eQTL gene mapping were primarily performed using the SNP2GENE function in FUMA. SNPs in high LD (r^2>0,6) with any independent lead SNPs (r^2 between independent SNPs <0.1) were mapped to genes if being physically located within or within 10kb of the genes, based on ANNOVAR annotations. FUMA also mapped SNPs to genes within a distance of 1Mb if associated with the expression of those genes in skin (reported in MuTHER or GTEx v7), dermal fibroblasts (GTEx v7 or v8 depending on SNP) and blood (reported in eQTLGen). A limited number of prioritised target genes were identified (Table 5.7): five genes for the 1q32.1 (*PHLDA3*, *NAV1*, *CSRP1*, *RP11-134G8*, *RPS10P7*), three genes for the 1q41 locus (*FAM177B*, *RP11-815M8*, *RP11-400N13*) and one gene for the 15q21.3 locus (*NEDD4*).

Table 5.7: Genes prioritised by positional mapping and eQTLs using FUMA.

Symbol	Locus	Type	posMapSNPs	MaxCADD	eqtlMapSNPs	eqtlMap_tissue
RP11-466L17.1	1q32.1	lincRNA	0	0.000	1	Blood
RP11-134G8.8	1q32.1	lincRNA	6	24.300	13	Fibroblasts; Skin
PHLDA3	1q32.1	protein_coding	8	24.300	11	Skin
CSRP1	1q32.1	protein_coding	8	7.614	11	Blood
RP11-134G8.7	1q32.1	antisense	0	0.000	11	Blood
RPS10P7	1q32.1	lincRNA	0	0.000	11	Blood; Skin
NAV1	1q32.1	protein_coding	0	0.000	11	Blood
RP11-815M8.1	1q41	lincRNA	0	0.000	1	Blood
RP11-400N13.3	1q41	lincRNA	1	0.572	23	Fibroblasts
RP11-400N13.1	1q41	lincRNA	123	15.160	0	—
QRSL1P2	1q41	pseudogene	2	0.297	0	—
FAM177B	1q41	protein_coding	0	0.000	13	Blood
NEDD4	15q21.3	protein_coding	26	12.270	20	Blood; Fibroblasts

eQTL, expression quantitative trait locus; SNP, single nucleotide polymorphism; posMapSNPs, number of SNPs mapped to gene based on positional mapping; MaxCADD, the maximum CADD score of mapped SNPs by positional mapping; eqtlMapSNPs, the number of SNPs mapped to the gene based on eQTL mapping; eqtlMap_tissue, tissue type of mapped eQTL SNPs.

Additional candidate eQTLs were identified by querying the eQTL catalogue of Phenoscanner v2 (a database of publicly available genotype associations with a broad range of phenotypes/traits)(Kamat et al., 2019). This yielded eQTLs associated with the expression of 10 genes in multiple tissues at P<1x10^{-5}(Table 5.8). This relaxed p-value was selected to maximise the number of genes for further analysis.

Table 5.8: Genes prioritised by eQTLs using Phenoscanner.

Symbol	Locus	Tissue	PMID
RP11-134G8.2	1q32.1	Muscle skeletal; Skin sun exposed lower leg	25954001
CSRP1	1q32.1	Whole blood	27918533; eQTLGen
NAV1	1q32.1	Whole blood	eQTLGen
RP11-134G8.7	1q32.1	Whole blood	eQTLGen
RPS10P7	1q32.1	Whole blood	eQTLGen
RPS10P7	1q32.1	Artery aorta; Esophagus muscularis	25954001
RPS10P7;RP11-134G8.6	1q32.1	Whole blood	27918533
RP11-400N13.3	1q41	Cells transformed fibroblasts	25954001
TAF1A	1q41	Colon transverse	25954001
NEDD4	15q21.3	Whole blood; Monocytes	28122634; 22446964
NEDD4	15q21.3	Whole blood	27918533; eQTLGen
NEDD4	15q21.3	Neutrophils; Artery tibial; Thyroid; Monocytes; Cells transformed fibroblasts; Colon transverse; Artery aorta; Atherosclerotic internal thoracic artery	27863251; 25954001; 25720628
PRTG	15q21.3	Artery aorta; Artery tibial; Atherosclerotic internal thoracic artery; Atherosclerotic aortic root	25954001; 25720628

eQTL, expression quantitative trait locus; PMID, PubMed identifier

The significant eQTLs (P<1x10^{-5}) were investigated further by using Approximate Bayes Factor colocalisation analysis(Giambartolomei and al, 2014). This quantifies the probability that expression of the analysed gene (in tissue or blood)

and an excessive scarring phenotype share the same causal variant by combining information across multiple variants within a genetic region (taking into account the prior probabilities that any random variant is associated with either or both traits). The analysis which was limited to the UKB study, showed no evidence for a single shared variant (PP.H4.abf > 0.75) with the expression of any of the selected genes (Table 5.9).

Table 5.9: Bayesian colocalisation to estimate posterior probability of single shared variant between the expression of mapped genes in skin, fibroblasts or blood and excessive scarring susceptibility.

Gene symbol	Locus	eQTL dataset	PP.H0	PP.H1	PP.H2	PP.H3	PP.H4
RP11-134G8.8	1q32.1	Skin GTEx v7	—	0.6798	—	0.2924	0.0278
PHLDA3	1q32.1	Skin GTEx v7	0.0092	0.6188	0.0039	0.2660	0.1021
CSRP1	1q32.1	eQTLGen	—	0.6944	—	0.2872	0.0184
RP11-134G8.7	1q32.1	eQTLGen	—	0.6921	—	0.2778	0.0301
RPS10P7	1q32.1	eQTLGen	—	0.6987	—	0.2829	0.0184
NAV1	1q32.1	eQTLGen	—	0.7197	—	0.2618	0.0185
RP11-815M8.1	1q41	eQTLGen	—	—	—	1.0000	—
RP11-400N13.3	1q41	Fibroblast GTEx v8	—	—	—	1.0000	—
FAM177B	1q41	eQTLGen	—	—	—	1.0000	—
NEDD4	15q21.3	Fibroblast GTEx v7	0.0002	0.3059	0.0002	0.4682	0.2255
NEDD4	15q21.3	Skin GTEx v7	0.3212	0.1013	0.4145	0.1306	0.0324
NEDD4	15q21.3	eQTLGen	—	0.4018	—	0.4906	0.1076

eQTL, expression quantitative trait locus; PP, posterior probability. For each row, probability values in columns PP.H0 to PP.H4 sum to 1; PP.H0 is the probability that neither trait has a genetic association in the region; PP.H1 indicates that only the eQTL has a genetic association in the region; PP.H2 indicates that only excessive scarring has a genetic association in the region; PP.H3 indicates that both traits are associated, but with different causal variants; PP.H4 indicates that both traits are associated and share a single causal variant.

5.4.3 Gene-level association analysis

Gene-level association analysis of the meta-analysis using MAGMA (a tool for gene-based association analysis which links variants to genes and integrates GWAS data to score genes for association with the phenotype as detailed in Chapter 2, embedded within FUMA) mapped all SNPs to 18831 protein-coding genes. Genome-wide significance was defined at P=0.05/18831=2.655x10^{-6}. Using the default SNP-wise (mean) model for the gene analysis, the *PHLDA3* association reached the lowest p-value of 3.53x10^{-8}. MAGMA gene-set analysis did not identify any significant associations with any specific groups of genes sharing biological or functional characteristics.

5.4.4 Phenotype associations

Phenoscanner was also used for look-up of published phenotype associations for the meta-analysis 95% credible set variants. Five associated traits were identified with P<10^{-5}: breast cancer, Dupuytrens disease, fibroblastic disorders, hair balding pattern 4 and sitting height; notably, the risk allele for Dupuytrens disease was directionally consistent (Table 5.10).

Table 5.10: Phenotype associations of the meta-analysis 95% credible set variants.

				Phenoscanner traits			Keloid (UKB)		Keloid (FG)		Keloid (meta-analysis)	
SNP	**Locus**	**Ref**	**Alt**	**Trait**	**OR (95%CI)**	**P**	**OR (95%CI)**	**P**	**OR (95%CI)**	**P**	**OR (95%CI)**	**P**
rs35383942	1q32.1	C	T	Breast cancer[a]	0.89 (0.87; 0.92)	4.00e-13	1.35 (1.13; 1.63)	1.04e-03	1.53 (1.32; 1.78)	9.83e-09	1.46 (1.3; 1.64)	6.99e-11
rs35383942	1q32.1	C	T	Hair or balding pattern: pattern 4[b]	0.98 (0.98; 0.99)	3.00e-08	1.35 (1.13; 1.63)	1.04e-03	1.53 (1.32; 1.78)	9.83e-09	1.46 (1.3; 1.64)	6.99e-11
rs1509406	15q21.3	A	G	Dupuytrens disease[c]	1.16 (1.11; 1.22)	3.00e-10	1.18 (1.07; 1.31)	9.90e-04	1.21 (1.1; 1.34)	1.09e-04	1.2 (1.12; 1.29)	3.98e-07
rs11632096	15q21.3	A	G	Fibroblastic disorders[b#]	1 (1; 1)	4.31e-07	1.2 (1.09; 1.33)	2.88e-04	1.22 (1.1; 1.35)	7.47e-05	1.21 (1.13; 1.3)	8.04e-08
rs28773028	15q21.3	C	G	Sitting height[b]	0.99 (0.99; 0.99)	4.57e-07	1.19 (1.07; 1.32)	7.06e-04	1.22 (1.1; 1.35)	1.01e-04	1.2 (1.12; 1.29)	2.63e-07

[a]PMID 29059683; [b]UKB(http://www.nealelab.is/uk-biobank/); [c]PMID 28886342

[#]Fibroblastic disorders is a group of disorders that encompass Dupuytrens disease, knuckle pads and unspecified fibrosis.

UKB, UK Biobank; FG, Finngen; Ref, reference allele; Alt, alternate allele; OR, odds ratio; 95%CI, 95% confidence interval

5.5 Subgroup analysis - black British participants

Variants around the three independent loci were also analysed for associations with excessive scarring in a subset of black British participants with relatively homogeneous genetic ancestry (n cases = 50, n controls =1501). As rs35383942 was filtered out during genotype QC for this sub-cohort, the next most significant SNP at the same locus, rs12134525, was used as a proxy in the analysis. Using a Bonferroni-corrected significance threshold of p<0.05/3=0.0125, no significant shared associations were found (Table 5.11).

Table 5.11: Association analyses for the three lead excessive scarring risk variants in white and black British participants.

SNP	Locus	Ref	Alt	Ethnicity	EAF	OR (95% CI)	P
rs12134525	1q32.1	T	C	White	0.93	0.72 (0.6; 0.86)	2.09e-04
rs12134525	1q32.1	T	C	Black	0.97	1.52 (0.41; 5.58)	5.13e-01
rs10863683	1q41	G	C	White	0.33	1.48 (1.34; 1.63)	1.32e-15
rs10863683	1q41	G	C	Black	0.25	0.93 (0.57; 1.54)	7.94e-01
rs60890210	15q21.3	A	G	White	0.73	0.82 (0.74; 0.9)	1.55e-04
rs60890210	15q21.3	A	G	Black	0.72	0.79 (0.49; 1.27)	3.16e-01

Ref, reference (non-effect) allele; Alt, alternate (effect) allele; EAF, effect allele frequency; OR(95%CI), odds ratio (95% confidence interval)

5.6 Discussion

By utilising primary and secondary clinical coding data in the UKB, SNPs at genome-wide significance levels associated with excessive scarring were identified in individuals of white British ancestry (843 cases and 187812 controls). Meta-analysis of the UKB dataset with the FinnGen GWAS of 293015 individuals (1126 cases and 291889 controls) identified independent SNPs in three loci to be associated with excessive scarring: 1q32.1 ($p=6.991x10^{-11}$), 1q41 ($p=3.034x10^{-25}$) and 15q21.3

($p=6.723x10^{-08}$), replicating findings previously only identified in East Asian participants.

It is notable that the most significant locus in this study, 1q41, a region without a known protein-coding gene, was also found to be the most significant in previous Japanese and Chinese GWAS studies(Ishigaki et al., 2020; Nakashima et al., 2010; Zhu et al., 2013). The closest gene is a recently-identified long intervening non-coding RNA (lincRNA), *RP11-400N13.1*.

Long non-coding RNAs are a heterogeneous class of non-coding RNAs longer than 200 nucleotides (of which lincRNA is the largest subtype) that have no apparent protein-coding functions. They tend to be lowly expressed and lack evolutionary conservation, but show relatively cell/tissue-specific expression.As a group, they represent a distinct class of regulatory RNAs that modulate transcriptional activities(Yao et al., 2019), but to date, few lincRNAs have experimentally verified functions or disease associations(Zhang et al., 2021). Regarding *RP11-400N13.1*, little is known, although a genome-wide eQTL association study reported strong association between the cis-eQTLs (expressed tissue unclear) and the keloid-associated SNP (rs873549) previously highlighted in the Japanese GWAS(McDowell et al., 2016; Nakashima et al., 2010).

Relatively more is reported regarding its next closest gene, *RP11-400N13.3* (*LINC01705*, ENSG00000232679, ENST00000438158.1), which is enriched in cultured fibroblasts (GTEx Portal v8). It was first identified as a differentially-upregulated lincRNA in gastric, colorectal and breast carcinomas and predicted to regulate ECM organisation(Bradford et al., 2016; Wang and Zhang, 2018). More recently, it was reported to be upregulated in retinal pigment epithelial (RPE) cells after stimulation with TGFβ1(Yang et al., 2021). A search of the literature yielded three functional studies for *RP11-400N13.3*; two studies found that it mediates TGFβ1 effects (in human lung fibroblasts(X. Yang et al., 2020) and in RPE cells(Yang et al., 2021)) with resultant upregulation of ECM markers and contractility(Yang et al., 2021). A separate study in colorectal cancer reported that it targets the miR4722-3p/P2RY8 miRNA/protein axis and promotes cellular proliferation, migration and invasion(H.

Yang et al., 2020). It is interesting however that the excessive scarring risk allele (rs10863683, C allele) is associated with a decreased expression of *RP11-400N13.3* in dermal fibroblasts (Table 5.12). Future functional study of *RP11-400N13.3* in dermal fibroblasts should provide clarification on its relevance to fibrosis.

Table 5.12: Fibroblast eQTL data restricted to rs10863683 for *RP11-400N13.3* in dermal fibroblasts (GTEx v8).

Gene	Variant	Ref	Alt	MAF	OR (95% CI)	P
RP11-400N13.3	rs10863683	C	G	0.372671	1.33 (1.18; 1.49)	1.51e-06

eQTL, expression quantitative trait locus; Ref, reference allele; Alt, alternate allele; MAF, minor allele frequency; OR(95%CI), odds ratio (95% confidence interval)

The second most significant locus, 1q32.1, was associated with several genes, with *PHLDA3* (Pleckstrin Homology Like Domain Family A Member 3) being the most biologically promising.

PHLDA3 was also prioritized in a recent large-scale Japanese GWAS of keloids[†], when it was mapped to a different lead SNP, rs192314256, with a large effect size (OR = 9.56, 5.91-15.45, p=3.28x10^{-20})(Ishigaki et al., 2020). Rs192314256 is rare with a MAF of 0.0096 in East Asians and 0 in Europeans (based on the ALFA project which provides aggregate allele frequency from dbGAP(Phan et al., 2020)); indeed this variant was not present in our meta-analysis. On the other hand, rs35383942, the putative causal SNP at this locus in the meta-analysis has a MAF of 0.058 in Europeans and is considerably rarer in East Asians (MAF 0.0002). Suitable proxy SNPs were not identifiable for comparison of direction of effect; nevertheless the candidate causal gene for this locus, *PHLDA3* is an attractive one as it is a repressor of AKT(Chen and Ohki, 2020; Kawase et al., 2009), whose signalling pathway

[†]An ambiguous definition for keloid that potentially includes hypertrophic scar cases was used in this study.

is known for its role in cell survival(Nicholson and Anderson, 2002), fibroblast activation(Li et al., 2016) and collagen production(Bujor et al., 2008).

The 15q21.3 locus identified in the meta-analysis[‡] overlaps with *NEDD4*, the first gene suggested in a GWAS to be associated with excessive scarring(Nakashima et al., 2010). *NEDD4* encodes a ubiquitin-protein ligase with a broad expression pattern and diverse *in vivo* functions(Boase and Kumar, 2015). Of relevance to keloids include a role as a regulator of inflammation (increased NF-κB/STAT3-mediated inflammation)(Marneros, 2019) and a regulator of ECM production (increased insulin-like growth factor 1/TGFβ1 signalling)(Daian et al., 2003; Fukushima et al., 2015). Interestingly, many of the most significant SNPs in this locus were associated with another disease of local cutaneous fibroproliferative growths, namely Dupuytrens disease(Ng et al., 2017). Despite similar histopathology, Dupuytrens disease has distinct clinico-epidemiological features to keloids - the fibrosis is usually limited to hands and commonly affects white middle-aged men(Karbowiak et al., 2021). The simultaneous occurence of both diseases is rare but has been reported(Gonza´lez-Marti´nez et al., 1995; Ly and Winship, 2011; Tsekouras and McGeorge, 1999). This potential shared genetic risk strengthens the evidence for the pleiotropy of *NEDD4* across heterogeneous fibrotic diseases and merits its further study.

Although the aforementioned mapped genes are functionally promising in their relevance to excessive scarring, Bayesian colocalisation showed no shared causal variant affecting their tissue expression (blood, fibroblasts, skin) and excessive scarring susceptibility, potentially indicating that the risk is mediated through other genes or tissue types that were not studied. However, the absence of convincing colocalisation does not mean that changes in gene expression are not mediators of genetic effects on disease risk; it may be a reflection of reduced power from sample size limitations

[‡]The FinnGen study identified a risk locus at chromosome 15 (comprising 24 rare variants unique to FinnGen at 48.6-52.8Mb with MAF <0.1%) that is close to but does not correspond precisely to 15q21.3. The Finnish population underwent a founding bottleneck ~120 generations prevously (as a result of rapid population expansion), which has resulted in Finnish-enriched variants that are rare or missing in other European populations; it is thus likely these novel variant associations are Finnish population-specific. The biological relevance to the established association is unclear and difficult to pursue further without Finnish genotype data to disentangle potential patterns of LD.

of the datasets analysed(Guo et al., 2015). The eQTL datasets analysed are also only available in bulk tissues and are not sensitive to the relevant differences in specific cell types. Therefore it is possible that if disease is caused by altered gene expression in a small subpopulation of cells, posterior probabilities of colocalisation may not be sufficiently high(Censin et al., 2021). Furthermore, the expression of genes in cells or tissues at a single point in time/ at a single state of activation is not representative of all the potential contexts of fibrosis/wound healing. Nevertheless, increasing availability of samples and a wider variety of cellular and tissue states will allow a more complete assessment of whether specific gene expression affects genetic risk of disease(Guo et al., 2015). Ultimately, the integrative analyses provide direction for further study and experimental validation is required to investigate the relevance of the mapped genes.

To conclude, this chapter reports the association of three loci with excessive scarring for the first time in individuals with white ethnicity. Previously only reported in East Asians, these risk loci have now been independently validated, which adds confidence to their biological relevance. Interrogation of these loci illustrates plausible biological processes through which the genetic risk is mediated. However, through the integrative bioinformatics approaches taken, there was no robust evidence for gene-set enrichment or causal gene identification.

5.7 Limitations and future plans

There have not been formal estimates of the heritability of excessive scarring but numerous reports suggest a heritable component(Glass, 2017; Shih and Bayat, 2010). The low SNP-based heritability estimates for excessive scarring in this study precluded genetic correlation studies to search for evidence of shared genetic architecture with diseases that were found to be phenotypically associated with excessive scarring (4). It is possible that there is a substantial environmental component for excessive scarring risk, although the low heritability estimates likely reflect limited statistical power rather than true lack of heritability(Nealelab, 2017). Indeed, for

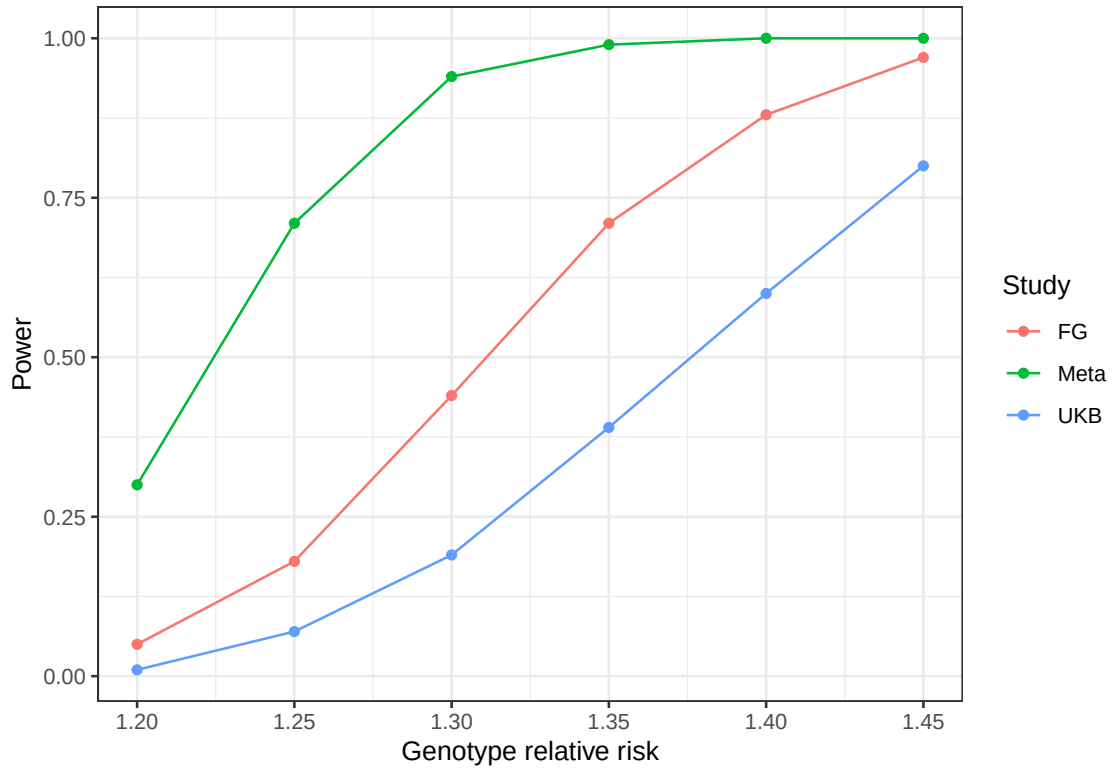

Figure 5.6: GWAS statistical power under the additive disease model based on different genotype relative risks. FG, Finngen; UKB, UK Biobank; Meta, meta-analysis; Power calculations performed using the GAS Power Calculator online tool (Johnson et al 2017)

the current study, the power to detect risk variants at genome-wide significance $p<5x10^{-8}$ with a risk allele frequency of 30% and a genotype relative risk <1.3 decreases substantially (Figure 5.6)(Johnson and Abecasis, 2017). The lack of evidence for association for the lead SNPs in black participants again reflects reduced statistical power from the small sample sizes; the power to detect the selected risk variants with allele frequencies of 30% and genotype relative risk of 1.5 at a Bonferroni-corrected p-value of 0.0125 was 45%(Johnson and Abecasis, 2017). Future plans to expand the GWAS by including data from other biobanks that have linkage to primary care (e.g. the Norwegian HUNT Biobank or the American NIH All of Us or Penn Medicine Biobanks) will boost statistical power to widen the detection of excessive scarring risk variants.

The use of clinical coding systems in case definitions may lead to phenotypic heterogeneity. As well as making it challenging to identify subtype-specific variants

(e.g. based on severity/clinical response or anatomical location), increasing the sample size of heterogeneous phenotypes may even increase the estimates of significance of genetic effects (i.e. p-values)(Kulminski et al., 2016). This presents a case for the recruitment of excessive scarring GWAS cohorts through specialist dermatology or plastic surgery centres, to allow the examination of more homogeneous subphenotypes, for the better understanding of their genetic influences.

6

Discussion and future directions

In the first half of this chapter, I discuss the key findings for the functional work related to *C1QTNF12* LOF and suggest avenues for its further study. This is followed by a discussion of the key findings from the observational population studies, with a focus on the PheWAS and GWAS studies. Limitations are acknowledged and future directions are suggested.

6.1 Homozygous *C1QTNF12* LOF variant in keloid pedigree

At the start of this thesis, I set out to investigate the hypothesis that adipolin LOF contributed to cutaneous fibroproliferation/fibrosis. The discovery of the homozygous frameshift LOF variant in *C1QTNF12* (encoding adipolin) which segregated with an unusually excessive scarring phenotype seemed compelling; it was the only LOF variant among just two homozygous protein-altering variants identified through WES - the other was a missense variant in *SEMG1* (rs7263910), encoding semenogelin-1, the predominant protein in semen, hence unlikely relevant to our skin phenotype. Analysis of the compound heterozygous variants shared by both siblings was also unremarkable. The *C1QTNF12* variant was validated by Sanger sequencing (both in blood and dermal fibroblasts, to account for somatic mosaicism).

However, confirmation of the predicted absent protein product was hampered by the lack of a positive control - adipolin was not detectable in the normal human dermis, subcutaneous fat or dermal fibroblast lines tested for this thesis.

Functional analysis showed lack of significant relevance for *C1QTNF12* in a fibrotic phenotype. Patient-derived *C1QTNF12*-LOF fibroblasts appeared more contractile in response to TGFβ1 (however WT fibroblasts do not appear to express adipolin) and treatment of fibroblasts with adipolin did not attenuate TGFβ1-induced contractility. These findings prompted a larger scale search for evidence supporting *C1QTNF12* pathogenicity (i.e. if disease-causing, this variant should either be present in other individuals with the same class of disorders or be absent in unaffected individuals(Narasimhan et al., 2016b)). The discovery of two unrelated individuals with linked primary and secondary care health records carrying the same homozygous variant (within the UKB 200K exomes) was opportune. Examination of their records showed no remarkable phenotype. Unfortunately, recall of these individuals for deep phenotyping was not possible. Notwithstanding other caveats that include poor genotype quality, potential variable penetrance and health recording quality, this was evidence *against* homozygous *C1QTNF12*-LOF being disease-causing.

Functionally-redundant homozygous gene knockouts exist. Several studies in large bottlenecked or consanguinous populations have confirmed the presence of genes with homozygous LOF variants in apparently healthy individuals(Lim et al., 2014; MacArthur et al., 2012; Narasimhan et al., 2016a; Rausell et al., 2020; Saleheen et al., 2017; Sulem et al., 2015). If the variants are truly LOF (rather than undergoing read-through or translation reinitiation), functional redundancy may exist because of alternative means (e.g. by paralogous genes). It has been shown that human monogenic diseases frequently have functionally-dispensible paralogs (e.g. through gene duplication leading to additional copies of genes that have similar functions), an evolutionary-beneficial phenomenon that is thought reduce the probability of disease genes being removed from the population(Chen et al., 2013). Although this would mask the phenotypic consequences in the majority of

individuals, it may be that in some subpopulations, this functional 'compensation' is lost (e.g. if functional duplicates are very lowly expressed), resulting in disease(Chen et al., 2013). Compensatory mechanisms may also exist that can recover the deficits from absent genes. For example, a substantial proportion of humans can survive with a complete deficiency of *TLR5*, which encodes for a cell surface receptor for bacterial flagellin(Casanova et al., 2011); alternative immune mechanisms provide sufficient protection in its absence(Zhao and Shao, 2015).

For many LOF variants, experimental characterisation is lacking - their interpretation tends to rely on indirect measures of functional impact, such as evolutionary conservation(Richards et al., 2015). In the analysis of confirmed gene knockouts, finding out whether a relevant phenotype exists is key; cases of disappointing molecular targets are not infrequent. For example, the findings that the enzymatic activity of lipoprotein-associated phospholipase A2 (encoded by *PLA2G7*) was associated with an increased risk for coronary heart disease led to numerous efforts in the development of small molecule inhibitors of this protein(Huang et al., 2019). However, its inefficacy in reducing disease risk was proven in randomised controlled trials(O'Donoghue et al., 2014; STABILITY Investigators et al., 2014). As expected, when carriers of the LOF variants in *PLA2G7* were analysed, no risk reduction was observed(Saleheen et al., 2017).

Despite the limitations, results from this work add insight to understanding the functional relevance of *C1QTNF12* LOF and highlights the need to leverage expression data, experimental models and functional genomics to study its effects. The effort to identify the function of every protein-coding gene is already being undertaken for researchers of the mouse genome, through the International Mouse Phenotyping Consortium, by systematically knocking out genes and performing standardised phenotyping analyses across biological systems(Meehan et al., 2017); incidentally, a review of this database did not reveal a relevant phenotype in *C1qtnf12*-knockout mice. Although substantially more challenging, the Human Knockout Project(Perdigoto, 2017), currently in its pilot phase, has similar aims, to develop a comprehensive catalogue of human LOF variants.

6.2 Unexplored approaches for characterising adipolin function in excessive scarring/keloids

6.2.1 Experimental approaches

The experimental work in this thesis has focused on the (presumed) heterogeneous population of primary dermal fibroblasts in a submerged monolayer culture model. Although there are benefits to studying primary cells in their native physiological environment, there is potential masking of subtle changes to individual cellular subpopulations, at heterogeneous states of differentiation/activation, that limit data interpretation(Lappalainen and MacArthur, 2021). Since the first scRNA study describing the transcriptionally and functionally heterogeneous nature of human dermal fibroblasts in 2018(Tabib et al., 2018), there have been an exponentially increasing number of studies to refine this heterogeneity further, with an aim to therapeutically manipulate targeted cells that are ultimately responsible for driving disease(Shaw and Rognoni, 2020). However, whether the heterogeneity is a function of cellular state in response to the local microenvironment or whether it represents distinct cell lineages is still unknown(desJardins-Park et al., 2020). It is worth noting that the exploration of several published single cell fibroblast datasets(Liu et al., 2022; Philippeos et al., 2018; Tabib et al., 2018) did not reveal significant differential expression for *C1QTNF12* (which was expressed at very low levels across the datasets) between the cellular clusters defined.

It remains to be explored whether adipolin plays a role in other cells types that are relevant to scarring. Adipolin has been shown to promote an anti-inflammatory response (reduced TNFα, IL-1β and MCP1 expression in response to lipopolysaccharide) in murine macrophages(Enomoto et al., 2011; Ogawa et al., 2019; Takikawa et al., 2020). Although primarily referred to as a fibrotic disorder, keloid/excessive scar development also depends on inflammation(Wang et al., 2020). Of the inflammatory cells involved, macrophages are a major player, with the alternatively activated M2 macrophages being prominent in later wound healing (tissue remodelling) stages(Jin et al., 2018). M2 macrophages have also been found at higher levels in unaffected

skin of people susceptible to hypertrophic scars(Butzelaar et al., 2016). Functionally, macrophages have been reported to promote differentiation of myofibroblasts(Shook et al., 2018) through the secretion of TGFβ1 and PDGF-CC(Glim et al., 2013; Hesketh et al., 2017). In light of these findings, the effect of adipolin on macrophages (and subsequent fibroblast responses) may be an avenue to pursue.

As adipolin is an adipokine, the relationship between (adipolin-secreting) adipocytes and scarring would also be interesting to explore. In murine wound healing, adipocytes have been observed to regenerate from myofibroblasts(Plikus et al., 2017), a phenomenon suggested to be potentially scar-reducing. Adipocytes have also been shown to reprogramme to myofibroblasts and it is increasingly established that adipose tissue can modulate fibrotic outcomes in numerous tissues(Sun et al., 2013). It is speculative to suggest that adipocytes that have downregulated adipogenic genes (e.g. *C1QTNF12*) would be more prone to profibrotic differentiation; the absence of adipolin detection in our limited subcutaneous fat lysate samples steered this project away from the study of adipocytes, but if adipocyte cultures are more freely available, this hypothesis may be interesting to explore.

There are various disease models for studying keloids as prototypic excessive scars. The 2D-fibroblast monolayer culture model, used in this thesis, is the most established. However, this limits investigation to a single cell type in an artificial environment. Co-culture systems add a layer of physiological complexity and have enabled the study of the paracrine effects of other cells, such as mesenchymal stem cells, on keloid fibroblasts(Arno et al., 2014; Fang et al., 2016; Kuwahara et al., 2016; Liu et al., 2018). An alternative approach to the adipolin-conditioned media culture system used in this thesis could have been the culturing of fibroblasts with the cellular source of adipolin, if known(Lebeko et al., 2019). Higher dimensional/3D organotypic models (e.g. organoid cultures(Lee et al., 2012), organotypic cultures and *ex vivo* explant cultures(Bagabir et al., 2012b)) aim for a more physiological arrangement of cells and a microenvironment that is more representative of tissue state. They also enable the study of other experimental readouts that are relevant to keloids, such as ECM composition and architecture. For example, it may be of

interest to investigate whether adipolin exists within the matrisome of scar ECM and/or whether its modulation alters ECM properties. An important limitation is that these models may exclude potentially relevant cell types and do not consider the spatio-temporal dynamics of scar formation. Animal models, although of value in the study of wound healing, do not exist for the study of keloids(Lebeko et al., 2019).

6.2.2 Observational (population level) approaches

As long as data sharing continues to progress unhindered, there will be increasingly comprehensive cohorts of deeply genotyped and phenotyped samples to catalogue variants, naturally occuring in the human population. However, how genetic variants result in associated phenotypes is complex - there are multiple cellular, developmental and environmental conditions, acting through and interacting with a wide range of pathways, that result in disease(Lappalainen and MacArthur, 2021).

In this thesis, the phenotypic consequences of adipolin LOF were interrogated by searching for carriers of homozygous LOF variants and reviewing their linked health records for potentially relevant traits - no relevant phenotype association was identified. The availability of a large deeply-phenotyped exome-sequenced cohort thorough the UKB enabled this effort. However, a major challenge was that the variants are extremely rare, allowing only a descriptive analysis. An alternative method to increase study power, gene-based burden testing, performs comparisons by aggregating all rare variants (summarising their cummulative effect) and testing their association with relevant phenotypic groups(Lee et al., 2014). This is based on the assumption that other variants in the same gene responsible for disease may have the same direction and magnitude of effect on the trait. Gene-based burden testing has been systemically performed for >400K individuals within the UKB(Karczewski et al., 2022); again, no robust phenotype association was identified for *C1QTNF12* LOF. However, this study did not include primary care data, a main source of excessive scarring diagnosis, so may present an opportunity for further analysis.

The adipolin discoveries made so far seem to parallel that of adiponectin, the adipokine whose insulin-sensitizing, anti-inflammatory, anti-atherogenic properties

have been widely demonstrated in animal models(Turer and Scherer, 2012). Its favourable effects have even recently been extended to skin fibrosis(Lakota et al., 2012; Luo et al., 2021; Marangoni et al., 2017) and it is conceivable that efforts will be made to use it as a therapeutic target(Marangoni et al., 2017). However, its human studies have painted a complicated picture; circulating adiponectin levels associate with systemic metabolite levels in individuals with good metabolic health, but have also been linked to higher mortality amongst groups with a high risk of metabolic disease(Sattar and Nelson, 2008). Furthermore, although genetic association studies indicate that variants associated with circulating adiponectin are also associated with cardiometabolic outcomes, Mendelian randomisation (MR) studies which test for the presence of causal relations have been negative, suggesting that its circulating level is likely to be an epiphenomenon of disease(Borges et al., 2017; Yaghootkar et al., 2013). Although there have been studies correlating circulating levels of adipolin with disease(Babapour et al., 2020; Fadaei et al., 2019; Kasabri et al., 2019; Tan et al., 2014a), as yet, there is no data available on its associated genetic variants. As integrated proteomics datasets continue to grow(Ferkingstad et al., 2021), should this data be made available in the future, a similar MR approach may be useful to test for causality between adipolin and disease in a large cohort.

Overall, the possibility that the homozygous *C1QTNF12* LOF variant identified through WES in our excessive scarring pedigree is not disease-causing highlights the challenges of gene identification. The lack of supporting functional evidence raised several considerations: an alternative coding gene variant is causal (WES filtering yielded variants in seven other potential causative genes that may be revisited: (i) a homozygous missense variant in *SEMG1* and six compound heterozygous missense variants in *CDCC88B, FAM47A, IGFN1, NUDT, SOX30, and ZNF790*); (ii) an alternative non-coding causal variant exists (whole genome sequencing of the pedigree, if available, would have been an option to explore this possibility); (iii) the *C1QTNF12* variant associates with a different phenotype that confers a risk of keloids (e.g. through biological pathway-rewiring(Hu et al., 2016)); (iv) the *C1QTNF12* variant associates with an unrelated phenotype and the excessive

scarring phenotype in the pedigree is a severe manifestation of the common complex keloid disorder with polygenic and environmental influences(Albert and Kruglyak, 2015; French and Edwards, 2020).

6.3 Phenotypic profile of individuals with excessive scarring

A review of the pedigree's clinical information did not point to any Mendelian syndromes for which keloid/excessive scarring is a known manifestation - the proband has a history of dermatofibromata, obesity (requiring bariatic surgery) and iron deficiency anaemia whereas her brother has a history of acne vulgaris. Both parents who are unaffected by keloid/excessive scarring, are overweight. Whether any of these diseases have a relevant association with keloids/excessive scarring is unknown. Reliable knowledge of disease comorbidities is important to understanding complex disease relationships. Epidemiologically robust disease-disease associations are more likely to show shared molecular-level mechanisms (thereby creating a higher level understanding of disease networks). For example, numerous studies of the phenotypic associations of asthma eventually led to discoveries of shared genetic architecture with immune, cardiometabolic and neurological/mental health disorders(Blair et al., 2013; Zhou et al., 2022; Zhu et al., 2019, 2018). More topical examples include studies of COVID-19 comorbidities (e.g. cardiovascular disease, diabetes, hepatitis, lung/kidney disease) eludicating common molecular pathways that advanced knowledge on the pathology of viral infection (e.g. immune signalling pathways and platelet biology(Dolan et al., 2020)). Such findings offer new avenues for functional validation, disease prevention and treatment.

The studies of keloid comorbidities to date have been focused on single disease pairs (in isolation), based on speculated biological or demographic similarities. For example, the notion that keloids are hyperproliferative led to a Taiwanese study reporting a higher risk for keloid individuals in the development of skin cancers(Y.-Y. Lu et al., 2021). Studies implicating Th2 responses and polymorphisms in *ADAM33*

(an endopeptidase associated with asthma and atopic eczema) in keloids, prompted studies in Korean(Kwon et al., 2021) and Taiwanese(Lu et al., 2018) populations that reported associations between keloids and atopic eczema. The association between keloids and 12 other fibrotic diseases (including pulmonary, renal, endomyocardial, hepatic and retroperitoneal fibrosis) was examined in a Taiwanese population; the only disease found to be associated was uterine leiomyoma(Sun et al., 2014). Obervations of a shared higher prevalence of uterine leiomyoma, hypertension, and vitamin D deficiency with keloids in black and Asian populations have also led to formal testing in small cohorts, with some conflicting findings(El Hadidi et al., 2021; Harmon et al., 2013; Rutherford and Glass, 2017; Woolery-Lloyd and Berman, 2002; Yu et al., 2013). Although insightful, these studies lack a comprehensive evaluation of systemic health associations for comparison.

The UKB study of keloid/excessive scarring comorbidities herein was undertaken with the primary aim of systematically documenting the comorbidities of keloid/excessive scarring (in Europeans, with comparisons across ethnic groups) and the secondary aim of generating a list of diseases from which a search for shared pathobiology/genetic architecture can be undertaken. Of the four most-studied comorbidities selected for specific analyses (hypertension, uterine leiomyoma, vitamin D deficiency, atopic eczema), only the association with atopic eczema was statistically significant (OR 1.68, $p<0.001$) after fully accounting for confounders.

As 94% of UKB participants are white British, this risk profile may not be representative of other ethnicities. Fully-adjusted specific analyses within ethnic groups revealed associations with hypertension and uterine leiomyoma which were unique to black participants and associations with vitamin D deficiency in Asian participants. This may be a result of ethnic variation in disease prevalence; however ethnicity (and Townsend Deprivation Index) had been adjusted for in the analyses. Although interpretation is limited by the small sample sizes, this finding suggests disparate comorbidities between ethnic groups and warrants validation/ further study into their risk determinants (whether biological, socioeconomic, cultural, environmental or behavioural).

Interrogating a broad spectrum of phenotypes in large cohorts (collected by physicians from multiple different disciplines), by way of a PheWAS, provides a comprehensive assessment of correlated diseases. Limitations include ascertainment bias, which in this study likely led to an over-representation of dermatological associations. Cummulative diagnosis codes assigned to each individual may introduce some misclassification; this may be reflected by the enrichment of "other hypertrophic and atrophic disorders of the skin" which may be a differential diagnosis for excessive scars. Also, interpretation of disease patterns is limited by the lack of temporal associations/disease trajectories; associations of excessive scarring with surgical (and dermal inflammatory diseases) may well arise from skin wounds rather than their shared molecular biology. Nevertheless, similar reports of non-genetic applications of PheWAS's are becoming more prevalent with the availability of large biobanks(Fromme et al., 2022).

Such published studies to date do not include primary care data, as resources to map them to phecodes are not readily available. In this study, primary care data were comprehensively mapped to phecodes, maximising case discovery for disease groups while reducing misclassification of diseases that do not tend to lead to hospitalisation. The PheWAS revealed a wide range of previously-unreported correlations across disease systems; musculoskeletal disease and pain symptoms were the most significant non-dermatological associations whereas abnormal weight gain, heart valve replacement and breast hypertrophy were the associations with the largest effect sizes (albeit with small sample sizes). Although the overall associations did not suggest a clear pattern, they highlight a higher health burden for individuals with excessive scarring, potentially indicating an increased risk of poorer health outcomes, and offer insights into potential shared pathobiology.

The lack of significant PheWAS associations with fibrotic diseases in other systems is noteworthy. For these diseases, the pathogenic cells, extracellular architecture, signalling pathways and mediators implicated are broadly generalisable(Ricard-Blum and Miele, 2020). It is therefore interesting why some fibrotic diseases were

less significantly associated than others (e.g. Positively associated: Dupuytren's contracture, OR=1.8, p=0.001; chronic non-alcoholic liver disease, OR=1.15, p=0.54; urethral stricture, OR= 1.99, p=0.004; pulmonary fibrosis, OR=2.03, p=0.07; peritoneal adhesions, OR 3.68, p=0.0001. Negatively associated: systemic sclerosis, OR=$3.04x10^{-6}$, p=0.97; liver cirrhosis, OR=0.89, p=0.87). Small sample sizes for some of these conditions likely meant insufficient power to detect significant associations, as signals are generally driven by diseases with the largest sample numbers. However these observation suggest there may be commonalities/differences between fibrotic diseases yet to be uncovered, for a more informed classification of this phenotype.

The comorbidity study in this thesis would ideally be expanded through a meta-analysis with other integrated data sources/international biobanks, which over time may overcome the issues of sample size and reduced ethnic representation. It is important however to be aware of the limitations that remain: the inherent limitations of the observational study design mean no causal inference can be made, clinical information may be erroneous/incomplete, and confounders may be unaccounted for. Furthermore, although the hope is that knowledge gained will be complementary to molecular level discoveries, high comorbidity between disease groups (e.g. depression and metabolic disorders) may not necessarily show corresponding genetic relationships(Bulik-Sullivan et al., 2015a).

6.4 Genetic architecture of excessive scarring

The widely-cited heritable nature of keloids(Chen et al., 2006a; Clark et al., 2009; Goeminne, 1968; Marneros et al., 2001; Omo-Dare, 1975; Ramakrishnan et al., 1974) provided justification for an exploration of its genetic architecture. Following the comprehensive phenotypic profiling of the excessive scarring (keloid/hypertrophic scar) cohort in the UKB, a GWAS was performed investigating corresponding genotype associations, using data from white British individuals within the UKB. Meta-analysing results from the UKB study with results from FinnGen (made up of a

distinct European cohort with a unique genetic makeup) led to the identification of three loci, previously only known to be associated with keloids(/excessive scarring) in East Asian populations. This indicates shared genetic components with biological functions that are generalisable across populations. Understanding the existence of this shared polygenic archictecture also informs the design of future larger studies in the discovery of novel susceptibility loci.

A striking finding from the keloid GWASs to date (including the work here) is the highly significant association with disease risk in 1q41, a non-coding region with limited structural and functional information. At this locus, the nearest genes were two related lincRNA, *RP11-400N13.1* and *RP11-400N13.3*; published research on these molecules suggested potential mechanistic relevance to fibrosis, such as in the regulation of ECM organisation in lung fibroblasts(X. Yang et al., 2020). eQTL data from dermal fibroblasts were leveraged to investigate whether the expression of either of these genes may be a molecular mediator linking genotype to phenotype. It was encouraging that the lead keloid risk variant was a significant eQTL for *RP11-400N13.3* in dermal fibroblasts ($p=1.51x10^{-6}$), however this could be a chance finding as eQTLs are abundant(Liu et al., 2019). SNP-level colocalisation, a method that analyses the co-occuring patterns between eQTLs and GWAS signals, with the hypothesis that a causal variant contributes to both gene expression changes and disease susceptibility(Li and Ritchie, 2021), suggested no shared causal variant between keloid risk and *RP11-400N13.3* expression in dermal fibroblasts. This approach had several caveats: the single causal variant assumption does not consider the fact multiple causal variants may exist in proximity (although Bayesian fine mapping of causal variants for keloid risk suggested a single causal variant at this locus, making multiple causal variants less likely); the datasets used may have been underpowered as well as not representative of the cellular/tissue/temporal expression pattern of the candidate genes that are relevant during disease development (the specific cellular type/state/activity/spatio-temporal context in which risk alleles/candidate genes act to modify keloid risk is unclear); the overall lower tissue expression levels of lincRNA may have added statistical uncertainty (datasets with deeper RNA

sequencing may be required). Furthermore, eQTL analyses performed do not test the effects of causal variants on other molecular traits (e.g. genomic regulatory elements) that have less data availability than gene expression profiles(Cano-Gamez and Trynka, 2020). Nevertheless, although analysis of the 'intermediate' non-coding genome is inherently more challenging/less interpretable, the significance of the association suggests a need to understand its target interactome in the context of excessive scarring.

Gene-based association analysis for the meta-analysis results using MAGMA limited the mapped genes to those that were protein-coding(Sey et al., 2020). *PHLDA3* was nominated; its significance to keloid pathology has recently been alluded to(Ishigaki et al., 2020). The lack of other significantly associated genes (e.g. *NEDD4*, p=7x10^{-4}) is likely a result of an inadequate sample size. Association analyses at the level of the gene (rather than single variants) are arguably more functionally relevant and less susceptible to erroneous findings due to differences in allele frequency or LD structure between populations. However this requires the knowledge of all variations within/around the genes and their regulatory region for reliable association with the phenotype(Neale and Sham, 2004). MAGMA assigns SNPs to nearest genes but does not take into account functional genomics (e.g. long range regulatory interactions or tissue-specific regulatory relationships). Modifications to MAGMA have been made to address this(Sey et al., 2020; A. Yang et al., 2020); these may be explored in future studies, alongside other complementary eQTL-based gene annotation tools, such as transcriptome-wide association studies (TWAS, which imputes genotype-gene expression relationship based on eQTL association statistics and tests associations between predicted gene expression and the phenotype/trait(Wainberg et al., 2019)).

Although widely described to have a heritable component, the heritability index for keloids has rarely been formally quantified. SNP heritability for keloids/excessive scarring in Japanese, estimated using LDSC, ranged from 0.28-0.36(Ishigaki et al., 2020). In comparison, the SNP heritability of keloids in this UKB study was considerably lower (0.0068), thereby precluding genetic correlation analyses. This suggests the contribution of lower frequency variants or an underpowered study(Nealelab,

2017). However, both studies shared most of the risk loci and had a similar sample size. The difference in heritability therefore may be a reflection of different (or population-specific) variant effect sizes and the moderation of heritability by other factors (e.g. gene-environment/ gene-gene interactions)(Shi et al., 2021). Models that account for population-specific effect sizes are being developed(Shi et al., 2021), and may be worth exploring in the future.

Heterogeneity in cohort definitions in this and other keloid population studies (by combining diagnosis codes for keloids and hypertrophic scars) may have introduced bias in effect size estimates even though most cases that were included were strictly keloid. The rationale for the keloid (ambiguous)/excessive scar definition in this study was two-fold: to increase study power (lack of case stratification being a trade-off for larger sample size) and to reduce misclassification bias (e.g. misclassifying hypertrophic scar cases as 'normal scar' controls). Furthermore, other large scale biobank studies on keloids have used similar broad definitions. Sensitivity analysis using strict keloid definitions and excluding hypertrophic scar diagnoses from the control cohort showed consistent findings. This promising finding seems to suggest that there may be a shared 'fibrotic' genetic architecture between the excessive scar types.

The prospect of a 'fibrotic' genetic architecture raises the question as to whether it is feasible to perform a cross-trait GWAS/meta-analysis to identify genetic associations across mechanistically-relevant diseases (e.g. identified through analysis of keloid comorbidities). As well as having increased statistical power, the risk variants identified across the multiple traits can be integrated for genetic prediction(Maier et al., 2018). This approach, which is particularly useful for less prevalent diseases such as keloids, has been adopted in numerous studies, grouping by pathomechanism(Ferreira et al., 2017; Sakaue et al., 2021; Shirai et al., 2022) and organ system(Masuda et al., 2019), and has revealed novel insights into disease pathogenicity. Reliable knowledge of disease correlations, ideally strengthened by significant genetic correlation data, is crucial. In this thesis, potentially relevant keloid/excessive scar disease associations may have been masked by the stringent phenome-wide

significance threshold adopted (e.g. peritoneal adhesions) or not captured due to clinical coding deficiencies (e.g. post-surgical stenosis). Establishing carefully-curated comorbidities for targeted association studies in large-scale cohorts may be of value.

So far the contribution of low frequency and rare alleles (MAF $< 1\%$) remains unexplored. This group of variants is thought to potentially be more easily interpretable than the common variants discovered by GWASs, assuming that a variant strongly affecting disease risk is kept at a low frequency by negative selection and that these variants have occured more recently in human history therefore have fewer other variants in LD(Sazonovs and Barrett, 2018). However, studying rare variants requires more complex study design and association-testing considerations. The standard methods of genotyping and genotype imputation have decreasing accuracy with decreasing MAF. WES is less suitable for traits where known variation is in non-coding regions, whereas whole genome sequencing may have prohibitively high costs and computational burden. Standard GWAS statistical tests for single rare variants require either enormous sample or effect sizes. Variant aggregation methods (burden testing) circumvent the issue of reduced study power but they are limited by a lack of clear functional groupings for non-coding variants(Lee et al., 2014; Sazonovs and Barrett, 2018). It may be that rare variant testing is best targeted for cohorts that have a higher probability of carrying low frequency variants (e.g. unique keloid pedigrees(Marneros et al., 2004) or individuals with known disease status but low PRS(Sazonovs and Barrett, 2018)).

The proposed approaches to further the genetic studies, for example, by increasing sample sizes or by studying rare variant association may still be hampered by limited information on causality. MR is an increasingly used approach to infer causality from genetic association studies, whereby genetic variants are used as proxy measures for clinical interventions and used to test for associations between disease exposures and outcomes, on the premise that the proxy variants are associated with the exposure and not a confounder and that they influence the oucome only through the exposure(VanderWeele et al., 2014). This depends on having adequate study power to identify suitable variants and may not eliminate some issues with

confounding (e.g. by population structure, LD, genetic pleiotropy). Causation analysis of genomic studies is an active area of research with continuing development of methodological advancements(Hu et al., 2018), but ultimately, larger cohorts are necessary for their application.

6.5 Biobanks for dermatological research

Much of this thesis been made possible by the availability of UKB data. This biobanking effort has resulted in a high dimensional data resource with large sample sizes, encompassing longitudinal health-related information, biochemical and molecular data for diverse clinical, phenomic and genomic research. In this thesis, a number of challenges were apparent, related to electronic health record (EHR) phenotyping and relevent complementary data.

EHR phenotyping is typically automated using structured EHR data (clinical codes). There are issues with the accuracy, standardisation and completeness of these records(Denny, 2012). For example, EHR data tend to be fragmented (whether due to clinician perceptions on what is clinically relevant or disconnected health providers); coding systems change over time resulting in the need for mapping strategies to combine codes from different systems (such as those employed in this thesis); coding systems have variable specificity across phenotypes and may be based on clinical suspicion rather than confirmation of diagnosis(Denny, 2012); codes based on clinical suspicion may accumulate in records over time despite the assignment of alternative confirmed diagnoses. Furthermore, an EHR-linked biobank such as the UKB has de-identified data which prevents patients from being recontacted for validation. The gold standard in validating clinical codes is a review of clinical notes(Newton et al., 2013). Increasingly, machine-learning tools are being developed to leverage this information for better phenotyping performance but, so far, there are no unified implementation approaches(Carter et al., 2022).

For most dermatological conditions, current EHR data are insufficient for accurate diagnosis prediction and disease stratification - key factors for successful

phenomic/genomic research. Complementary data to validate EHR (e.g. clinical photographs/histology) are lacking in biobank studies. The most recent revision of the ICD codes, ICD11, substantially improves the coverage and categories of dermatological diseases - a result of the involvement of the international dermatology community. However, its timely implementation is unlikely and it may still lack structured terminologies of sufficient granularity to accurately capture dermatological presentations. Several initiatives were set up to create dermatology-specific EHR (e.g. DermLex(Papier et al., 2004) and DermO(Fisher et al., 2016)) but uptake has been limited and maintenance for both appear to have been discontinued. This indicates the need for greater dermatological presence within existing biobank studies to advocate for the integration of dermatology-relevant data. There is also the case for setting up more disease-relevant consortiums, especially for under-represented diseases such as keloids/hypertrophic scars.

6.6 Future directions

The *in vitro* and observational population data from this thesis do not support the hypothesis that homozygous *C1QTNF12* LOF, identified following exome sequencing of a small pedigree with excessive scarring, causes the observed phenotype. In light of its current little-known/unvalidated protein expression status, a data-driven approach will ideally be pursued further to identify supportive evidence, before *in vitro* functional studies. Examination of the 200K exomes in the UKB leaves the possibility that significant clinical phenotypes associated with *C1QTNF12* LOF were missed as a result of small sample sizes (with the identification of only two individuals with homozygous *C1QTNF12* LOF variants). With the exomes of over 470000 UKB participants now available, phenotypes robustly associated with LOF and/or missense variants in *C1QTNF12* could be re-interrogated. Absence of phenotypically-significant associations may help redirect efforts towards functional analyses of other candidate genes.

The observational population studies in this project have been restricted to British and Finnish populations. It would be worthwhile to harness the growing number of biobanks worldwide to expand on this work. Of immediate relevance are the the US-based NIH All of Us and the PennMed Biobank (PMBB). The NIH All of Us programme has ~1500 cases of keloid amongst 227740 participants. Although a similar data browser is not publicly available for PMBB, it is made up of a substantially greater proportion of black participants (30% for PMBB as compared to ~1% for UKB); the likelihood of a significant keloid cohort would be high. Even with the limited power of this work, from a comorbidity perspective, data from this thesis provided preliminary supporting evidence for a funding application to study the mechanistic link between keloids, hypertension and uterine leiomyoma in black women. The increased number of black participants by the addition of these datasets would be especially valuable to validate these associations and may allow genetic correlation analysis. From a genetic perspective, the prospect of potentially doubling the number of cases means greater power to detect less common risk variants as well as those with smaller effect sizes. Both biobanks however are currently only accessible by registered US-based institutions, therefore, collaborative links would need to be established - the Keloid Research Foundation(Uitto and Tirgan, 2020) may be a helpful springboard.

Other valuable biobank resources include the well-established Biobank Japan(Nagai et al., 2017) (which comprises the cohort for the first published keloid GWAS(Nakashima et al., 2010)) and Taiwan Biobank (from which keloid studies have not yet emerged)(Lin et al., 2020). Sakaue et al(Sakaue et al., 2021) performed cross-population meta-analyses for a large range of phenotypes, combining GWAS data from Biobank Japan, UKB and FinnGen; although the keloid/excessive scarring phenotype was included, only 668 cases were detected from UKB and FinnGen combined, compared with 1126 cases detected in this thesis. The Sakaue et al study(Sakaue et al., 2021), which identified a novel keloid variant, rs74983791*, used only a single ICD10 code

*This variant was only genome-wide significant for the meta-analysis, but not for the individual datasets analysed.

(L91.0) to define keloids (and hypertrophic scarring). Employing the manually-curated clinical codes used in this thesis and expanding the data sources further, will hopefully maximise study power whilst minimising misclassification.

Increasing the scale of a single type of -omic data is insufficient to fully explain genetic susceptibility. An example is the highly significant 1q41 keloid risk locus which contains a *cis*-eQTL for the lincRNA, *RP11-400N13.3* but subsequent colocalisation analysis failed to support the hypothesis that there was a shared causal variant. Although it is tempting to suggest *in vitro* functional analysis of *RP11-400N13.3*, manipulation of lincRNA is fraught with challenges(Gao et al., 2020) and numerous studies have reported ambiguous or insignificant resulting phenotypes(Amândio et al., 2016; Gao et al., 2020; Goudarzi et al., 2019; Han et al., 2018; Kölling et al., 2018; Stafford et al., 2016). It is pertinent to further explore bioinformatics evidence supporting causality (e.g. *in silico* prediction of its structural motifs/functional domains, correlation of its expression with CpG-site methylation and characterising its mRNA interaction network(Giral et al., 2018; Tonmoy et al., 2022)). Notably, these analyses still leave gaps in the understanding of the 1q41 locus (e.g. *trans*-eQTLs are not considered). Ultimately, the aim is to integrate multi-omic data to correlate genetic findings with protein and functional spatio-temporal localisation, in nominating targets for experimental validation. This should be performed in suitable *in vitro/in vivo* models and challenged with disease-relevant stimuli.

6.7 Conclusions

The identification of the homozygous *C1QTNF12* LOF variant in a single family with an autosomal recessive inheritence pattern of excessive scarring has led to a scientific exploration studying the role of adipolin in fibrosis, with the use of laboratory-based and bioinformatics approaches. The lack of strong functional evidence to suggest pathophysiological relevance led to data-driven studies of excessive scarring susceptibilities at a population level from a phenotypic and genetic perspective; these have

painted a multi-morbid/polygenic picture underlying the disorder whilst highlighting the need for better representation of skin phenotypes in EHR for future research.

Supplementary materials

Appendices

A
Supplementary data

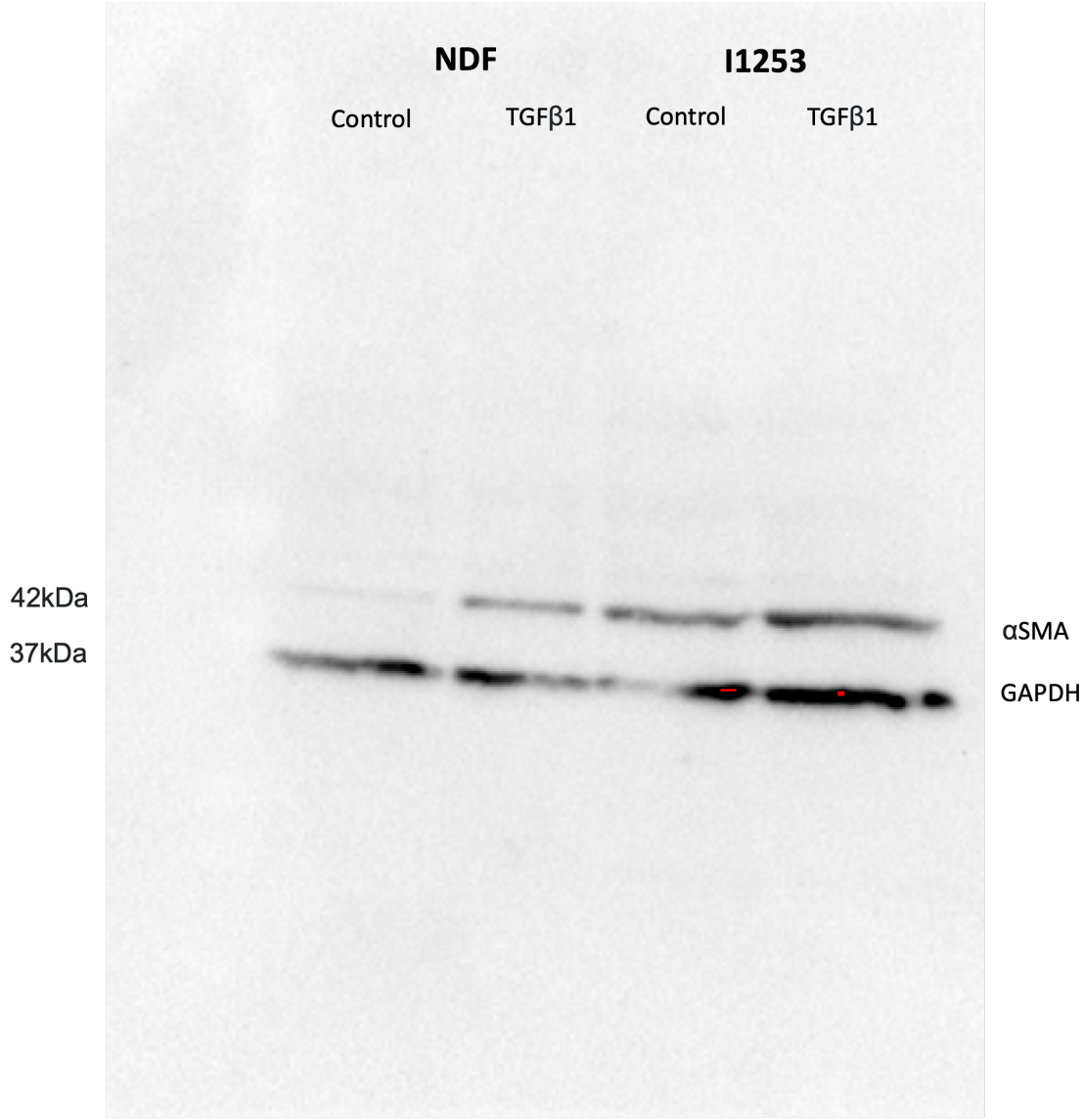

Figure A.1: Full Western blot image of NDF and I1253 lysates probed for alpha smooth muscle actin and GAPDH.

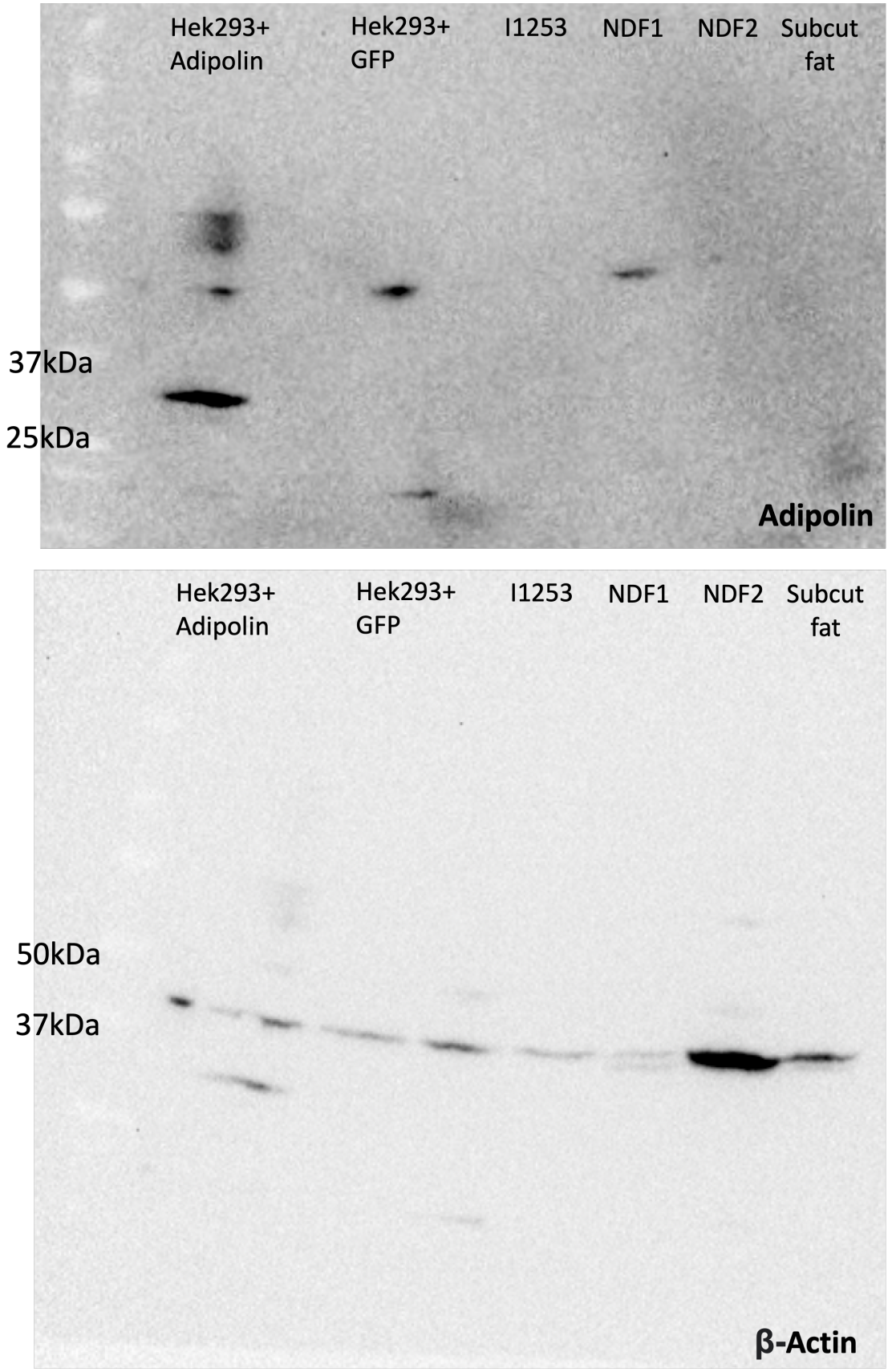

Figure A.2: Full Western blot image of protein lysates probed for adipolin and beta actin.

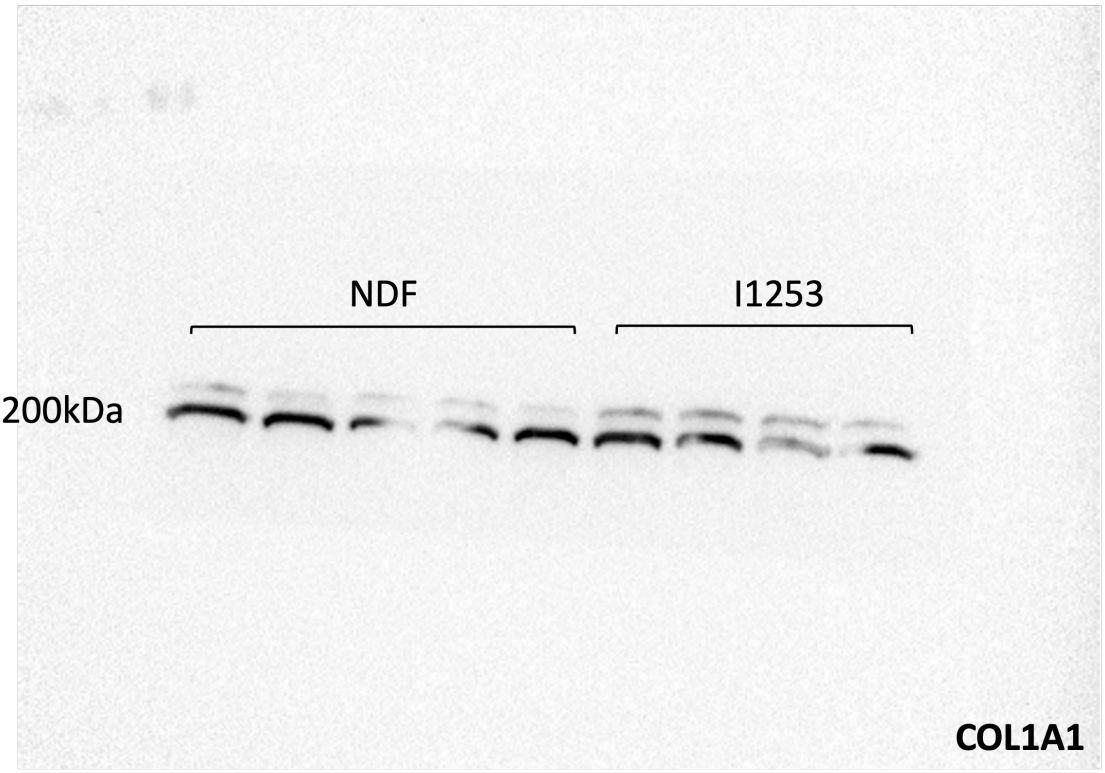

Figure A.3: Full Western blot image of protein lysates probed for collagen 1.

Table A.1: Clinical codelists for comorbidities in specific multivariable analyses.

Disease	ICD9	ICD10	Read2	Read3
Hypertension	4010, 4011, 4019	I10, I11, I12, I13, I15	14A2., 21261, 212K., 61462, 66, 662b., 662c., 662d., 662F., 662G., 662O., 662r., 7Q01., 8B26., 8BL0., 8I3N., 9OI9., F4042, F4213, G2..., G20.., G200., G201., G202., G203., G20z., G21.., G210., G2100, G2101, G211., G2110, G2111, G21z., G21z0, G21z1, G21zz, G22.., G220., G221., G222., G22z., G23.., G230., G231., G232., G233., G234., G23z., G24.., G240., G2400, G240z, G241., G2410, G241z, G244., G24z., G24z0, G24z1, G24zz, G2y.., G2z.., G672., Gyu2., Gyu21, L122., L1220, L1221, L1223, L122z, L127., L127z, L128., L1280, L1282, TJC7., TJC7z, U60C5	14A2., 21261, 9OI9., 61462, G24.., G240., G2400, G240z, G241., G2410, G241z, G244., G24z., G24z0, G24z1, Gyu21, 662F., 662G., F4211, G2..., G200., G201., G202., G20z., G21.., G210., G2100, G2101, G211., G2110, G2111, G21z., G21z1, G220., G221., G222., G22z., G23.., G230., G231., G232., G233., G234., G23z., G2y.., G2z.., Gyu2., L122., L1220, L1221, L1223, L122z, L127., L127z, L128., L1280, L1282, TJC52, TJC53, TJC7., TJC70, TJC71, TJC72, TJC73, TJC74, TJC75, TJC7z, U60C5, X00ef, XE0Ub, XE0Uc, XE0Ud, XE0Ue, XE0Uf, XE0Ug, XE0VM, XE15r, XM02V, XM1Qp, XSDSb, Xa8HD, XaBLq, XaIy8, XaIyC, XaIyD, XaIyE, XaJ5h, XaM5f, XaNFs
Leiomyoma of uterus	—	D25, D250, D251, D252, D259	7A54A, 7E061, 7E0DC, B78.., B780., B781., B782., B78z., BBK0., BBK00, BBK03, BBK05, BBK06, BBK0z	B78.., B780., B781., B782., B78z., BBK0., BBK00, BBK03, BBK05, BBK06, BBK0z, X403S, XE06e, Xa99u, XaNQc, XaQYZ

#Two self-reported codes were additionally retrieved for eczema, namely 1452 and 1669 (UK Biobank Data Coding 6).

Disease	ICD9	ICD10	Read2	Read3
Eczema[#]	512, 540, 3734A, 3735, 3735A, 691, 692, 693, 6984	B000, B001, B653, L20, L208, L209, L21, L211, L218, L219, L22, L23, L230, L231, L232, L233, L234, L235, L236, L237, L238, L239, L24, L240, L241, L242, L243, L244, L245, L246, L247, L248, L249, L25, L250, L251, L252, L253, L254, L255, L258, L259, L26, L27, L272, L278, L279, L30, L300, L303, L308, L309, L562, L58, L580, L581, L589, L710, L981	14F1., 26C4., A512., A5320, A540., A5441, A54x3, F4D30, F4D31, F4D4., F4D5., G831., G832., M101., M102., M11.., M110., M111., M112., M113., M114., M116., M117., M118., M1180, M118z, M119., M11A., M11z., M12.., M120., M121., M122., M1220, M1221, M1222, M1223, M1224, M1225, M122z, M123., M1230, M1231, M1232, M1233, M1234, M1235, M1236, M1237, M1238, M123z, M124., M1240, M1241, M1242, M1243, M1244, M1245, M1246, M1247, M1248, M124z, M125., M1250, M1251, M1252, M1253, M1254, M1255, M125z, M126., M1260, M1261, M1262, M1263, M1264, M1265, M1266, M126z, M127., M1270, M1273, M1278, M127z, M128., M1280, M1281, M1282, M1283, M1284, M1285, M1286, M129., M1290, M1291, M1292, M1293, M1294, M1295, M12C., M12C0, M12C1, M12y., M12y0, M12y1, M12y2, M12y3, M12y4, M12y5, M12y6, M12y7, M12y8, M12y9, M12yA, M12yB, M12yC, M12yD, M12yz, M12z., M12z0, M12z1, M12z2, M12z3, M12z4, M12zz, M13.., M130., M131., M13y., M13z., M140., M1535, M1536, M1831, M184., M1y0., M1y2., Myu2., Myu20, Myu21, Myu22, Myu23, Myu24, Myu25, Myu26, Myu27, Myu28, Myu29, Myu2C	14F1., A5320, A540., A5441, A54x3, F4D31, F4D4., F4D5., G831., G832., M07y., M1..., M100., M101., M102., M11.., M110., M111., M112., M113., M114., M117., M118., M118z, M11z., M12.., M120., M121., M122., M1220, M1221, M1222, M1223, M1224, M1225, M122z, M123., M1230, M1231, M1232, M1233, M1234, M1235, M1236, M1237, M1238, M123z, M124., M1240, M1241, M1242, M1243, M1244, M1245, M1246, M1247, M1248, M124z, M125., M1250, M1251, M1252, M1253, M1254, M1255, M125z, M126., M1260, M1261, M1262, M1263, M1264, M1265, M1266, M126z, M127., M1270, M1273, M1278, M127z, M1280, M1281, M1282, M1283, M1284, M1285, M1286, M1290, M1291, M1292, M1293, M1294, M12C., M12C0, M12C1, M12y., M12y0, M12y1, M12y2, M12y3, M12y4, M12y5, M12y6, M12y7, M12y8, M12y9, M12yA, M12yB, M12yC, M12yD, M12yz, M12z., M12z0, M12z1, M13.., M130., M131., M13y., M13z., M184., M2y41, Myu2., Myu20, Myu21, Myu23, Myu24, Myu25, Myu26, Myu27, Myu28, Myu29, Myu2C, Ua1AR, X00YT, X00YV, X00YW, X00iS, X30Cp, X40Fx, X503k, X503l, X503m, X5040, X504O, X504t, X504u, X504v, X5051, X505H, X505K, X505L, X505M, X505N, X505O, X505P, X505Q, X505R, X505S, X505T, X505U, X505V, X505X, X505Y, X505Z, X505a, X505b, X505c, X505d, X505e, X505f, X505i, X505j, X505k, X505l, X505o, X505p, X505r, X505s, X505t, X505u, X505v, X505w, X505x, X505z, X5060, X5061, X5062, X5063, X5064, X5065, X5066, X5067, X5068, X5069, X506C, X506J, X506K, X506L, X506O, X506P, X506Q, X506U, X506c, X506d, X5082, X50AK, X50Gj, X50Gk, X50J8, XE16i, XE1An, XE1Ap, XE1Aq, XE1Ar, XE1As, XE1At, XE1Au, XE1Av, XE1Aw, XE1C4, XE1C6, XE1C8, XE1CA, XE1CC, XE1CE, XE1CG, XE1CI, XE1CW, XM1PZ, XM1Pa, Xa0WJ, Xa0p8, Xa13g, Xa1dl, Xa3gI, Xa3gJ, Xa7lZ, Xa7lb, Xa9Bb, Xa9CV, XaBml, XaBmm, XaBsL, XaEJY, XaEJZ, XaINK, XaINM, XaL2Q, XaY4Z, XaY4a, XaY4o, XaYX9, .14F1, .26C4, .F5C4, .G932, .L2.., .L21., .L22., .L221, .L22Z, .L23., .L231, .L232, .L233, .L234, .L235, .L23Z, .L24., .L241, .L242, .L243, .L244, .L245, .L246, .L247, .L248, .L24Z, .L25., .L251, .L252, .L25Z, .L27., .L28., .L2Z., .L318, .L353, .L354, 26C4., F4D30, M116., M119., M11A., M128., M129., M1295, M12z2, M12z3, M12zz, M1535, M1536, M1y0., M1y2., M21A., Myu22, X504w, X75uT, Xa4jb
Vitamin D deficiency	268	E55, E559	C28..	C28.., XE112, .C415

[#]Two self-reported codes were additionally retrieved for eczema, namely 1452 and 1669 (UK Biobank Data Coding 6).

Table A.2: Example mapping for primary care and hospital records to the phecode for 'Uterine leiomyoma'.

Phecode	Phecode description	Code type	Code	Description	ICD10 equivalent	ICD10 description
218.1	Uterine leiomyoma	ICD10	D259	Leiomyoma of uterus, unspecified	—	—
218.1	Uterine leiomyoma	ICD10	D25	Leiomyoma of uterus	—	—
218.1	Uterine leiomyoma	ICD10	D250	Submucous leiomyoma of uterus	—	—
218.1	Uterine leiomyoma	ICD10	D251	Intramural leiomyoma of uterus	—	—
218.1	Uterine leiomyoma	ICD10	D252	Subserosal leiomyoma of uterus	—	—
218.1	Uterine leiomyoma	Read2	B78..	Uterine leiomyoma - fibroids	D259	Leiomyoma of uterus, unspecified
218.1	Uterine leiomyoma	Read2	B78z.	Uterine leiomyoma NOS	D259	Leiomyoma of uterus, unspecified
218.1	Uterine leiomyoma	Read2	B781.	Intramural uterine leiomyoma	D251	Intramural leiomyoma of uterus
218.1	Uterine leiomyoma	Read2	B780.	Submucous uterine leiomyoma	D250	Submucous leiomyoma of uterus
218.1	Uterine leiomyoma	Read2	B782.	Subserous uterine leiomyoma	D252	Subserosal leiomyoma of uterus
218.1	Uterine leiomyoma	Read3	B78..	Uterine fibroid	D259	Leiomyoma of uterus, unspecified
218.1	Uterine leiomyoma	Read3	B781.	Intramural uterine fibroid	D251	Intramural leiomyoma of uterus
218.1	Uterine leiomyoma	Read3	B78z.	Uterine leiomyoma NOS	D259	Leiomyoma of uterus, unspecified

Phecode	Phecode description	Code type	Code	Description	ICD10 equivalent	ICD10 description
218.1	Uterine leiomyoma	Read3	B782.	Subserous uterine fibroid	D252	Subserosal leiomyoma of uterus
218.1	Uterine leiomyoma	Read3	B780.	Submucous uterine fibroid	D250	Submucous leiomyoma of uterus
218.1	Uterine leiomyoma	Read3	X78XX	Uterine fibroid polyp	D259	Leiomyoma of uterus, unspecified

Bibliography

References

Abdou, A.G., Maraee, A.H., Abd-Elsattar Saif, H.F., 2014. Immunohistochemical Evaluation of COX-1 and COX-2 Expression in Keloid and Hypertrophic Scar. The American Journal of Dermatopathology 36, 311–317. https://doi.org/10.1097/dad.0b013e3182a27b83

Abercrombie, M., 1978. Fibroblasts. Journal of Clinical Pathology. Supplement (Royal College of Pathologists). 12, 1–6.

Adams, CD, 2021. Colocalization of ch16 eQTLs and longevity (online, accessed may 20 2021).

Adotama, P., Rutherford, A., Glass, D.A., 2015. Association of keloids with systemic medical conditions: a retrospective analysis. International Journal of Dermatology 55, e38–e40. https://doi.org/10.1111/ijd.12969

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., Sunyaev, S.R., 2010. A method and server for predicting damaging missense mutations. Nature Methods 7, 248–249. https://doi.org/10.1038/nmeth0410-248

Akasaka, Y., Fujita, K., Ishikawa, Y., Asuwo, N., Inuzuka, K., Ishihara, M., Ito, M., Masuda, T., Akishima, Y., Zhang, L., Ito, K., Ishii, T., 2001. Detection of apoptosis in keloids and a comparative study on apoptosis between keloids, hypertrophic scars, normal healed flat scars, and dermatofibroma. Wound Repair and Regeneration 9, 501–506. https://doi.org/10.1046/j.1524-475x.2001.00501.x

Ala-Kokko, L., Rintala, A., Savolainen, E.-R., 1987. Collagen Gene Expression in Keloids: Analysis of Collagen Metabolism and Type I. III, IV, and V Procollagen mRNAs in Keloid Tissue and Keloid Fibroblast Cultures. Journal of Investigative Dermatology 89, 238–244. https://doi.org/10.1111/1523-1747.ep12471056

Al-Attar, A., Mess, S., Thomassen, J.M., Kauffman, C.L., Davison, S.P., 2006. Keloid Pathogenesis and Treatment. Plastic and Reconstructive Surgery 117, 286–300. https://doi.org/10.1097/01.prs.0000195073.73580.46

Albert, F.W., Kruglyak, L., 2015. The role of regulatory variation in complex

traits and disease. Nature Reviews Genetics 16, 197–212. https://doi.org/10.1038/nrg3891

Alford, P.W., Nesmith, A.P., Seywerd, J.N., Grosberg, A., Parker, K.K., 2011. Vascular smooth muscle contractility depends on cell shape. Integrative Biology 3, 1063–1070. https://doi.org/10.1039/c1ib00061f

Alipoor, E., Salmani, M., Yaseri, M., Kolahdouz-Mohammadi, R., Esteghamati, A., Hosseinzadeh-Attar, M.J., 2019. Role of type 2 diabetes and hemodialysis in serum adipolin concentrations: A preliminary study. Hemodialysis International 23, 472–478. https://doi.org/10.1111/hdi.12787

Amândio, A.R., Necsulea, A., Joye, E., Mascrez, B., Duboule, D., 2016. Hotair Is Dispensible for Mouse Development. PLOS Genetics 12, e1006232. https://doi.org/10.1371/journal.pgen.1006232

Amini-Nik, S., Yousuf, Y., Jeschke, M.G., 2018. Scar management in burn injuries using drug delivery and molecular signaling: Current treatments and future directions. Advanced Drug Delivery Reviews 123, 135–154. https://doi.org/10.1016/j.addr.2017.07.017

Ammendola, M., Zuccalà, V., Patruno, R., Russo, E., Luposella, M., Amorosi, A., Vescio, G., Sammarco, G., Montemurro, S., De Sarro, G., Sacco, R., Ranieri, G., 2012. Tryptase-positive mast cells and angiogenesis in keloids: a new possible post-surgical target for prevention. Updates in Surgery 65, 53–57. https://doi.org/10.1007/s13304-012-0183-y

Arima, J., Huang, C., Rosner, B., Akaishi, S., Ogawa, R., 2015. Hypertension: a systemic key to understanding local keloid severity. Wound Repair and Regeneration 23, 213–221. https://doi.org/10.1111/wrr.12277

Arno, A.I., Amini-Nik, S., Blit, P.H., Al-Shehab, M., Belo, C., Herer, E., Jeschke, M.G., 2014. Effect of Human Wharton's Jelly Mesenchymal Stem Cell Paracrine Signaling on Keloid Fibroblasts. Stem Cells Translational Medicine 3, 299–307. https://doi.org/10.5966/sctm.2013-0120

Ashcroft, G.S., 1999. Bidirectional regulation of macrophage function by TGF-? Microbes and Infection 1, 1275–1282. https://doi.org/10.1016/s1286-4579(99)00257-9

Ashcroft, K.J., Syed, F., Bayat, A., 2013. Site-Specific Keloid Fibroblasts Alter the Behaviour of Normal Skin and Normal Scar Fibroblasts through Paracrine Signalling. PLoS ONE 8, e75600. https://doi.org/10.1371/journal.pone.0075600

Atiyeh, B.S., 2020. Nonsurgical Management of Hypertrophic Scars: Evidence-Based Therapies, Standard Practices, and Emerging Methods. Aesthetic Plastic

Surgery 44, 1320–1344. https://doi.org/10.1007/s00266-020-01820-0

Aulchenko, Y.S., Koning, D.-J. de, Haley, C., 2007. Genomewide Rapid Association Using Mixed Model and Regression: A Fast and Simple Method For Genomewide Pedigree-Based Quantitative Trait Loci Association Analysis. Genetics 177, 577–585. https://doi.org/10.1534/genetics.107.075614

Austin, P.C., 2011. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behavioral Research 46, 399–424. https://doi.org/10.1080/00273171.2011.568786

Auton, A., Abecasis, G.R., Altshuler, D.M., Durbin, R.M., Abecasis, G.R., Bentley, D.R., Chakravarti, A., Clark, A.G., Donnelly, P., Eichler, E.E., Flicek, P., Gabriel, S.B., Gibbs, R.A., Green, E.D., Hurles, M.E., Knoppers, B.M., Korbel, J.O., Lander, E.S., Lee, C., Lehrach, H., Mardis, E.R., Marth, G.T., McVean, G.A., Nickerson, D.A., Schmidt, J.P., Sherry, S.T., Wang, J., Wilson, R.K., Gibbs, R.A., Boerwinkle, E., Doddapaneni, H., Han, Y., Korchina, V., Kovar, C., Lee, S., Muzny, D., Reid, J.G., Zhu, Y., Wang, J., Chang, Y., Feng, Q., Fang, X., Guo, X., Jian, M., Jiang, H., Jin, X., Lan, T., Li, G., Li, J., Li, Y., Liu, S., Liu, X., Lu, Y., Ma, X., Tang, M., Wang, B., Wang, G., Wu, H., Wu, R., Xu, X., Yin, Y., Zhang, D., Zhang, W., Zhao, J., Zhao, M., Zheng, X., Lander, E.S., Altshuler, D.M., Gabriel, S.B., Gupta, N., Gharani, N., Toji, L.H., Gerry, N.P., Resch, A.M., Flicek, P., Barker, J., Clarke, L., Gil, L., Hunt, S.E., Kelman, G., Kulesha, E., Leinonen, R., McLaren, W.M., Radhakrishnan, R., Roa, A., Smirnov, D., Smith, R.E., Streeter, I., Thormann, A., Toneva, I., Vaughan, B., Zheng-Bradley, X., Bentley, D.R., Grocock, R., Humphray, S., James, T., Kingsbury, Z., Lehrach, H., Sudbrak, R., Albrecht, M.W., Amstislavskiy, V.S., Borodina, T.A., Lienhard, M., Mertes, F., Sultan, M., Timmermann, B., Yaspo, M.-L., Mardis, E.R., Wilson, R.K., Fulton, L., Fulton, R., Sherry, S.T., Ananiev, V., Belaia, Z., Beloslyudtsev, D., Bouk, N., Chen, C., Church, D., Cohen, R., Cook, C., Garner, J., Hefferon, T., Kimelman, M., Liu, C., Lopez, J., Meric, P., O'Sullivan, C., Ostapchuk, Y., Phan, L., Ponomarov, S., Schneider, V., Shekhtman, E., Sirotkin, K., Slotta, D., Zhang, H., McVean, G.A., Durbin, R.M., Balasubramaniam, S., Burton, J., Danecek, P., Keane, T.M., Kolb-Kokocinski, A., McCarthy, S., Stalker, J., Quail, M., Schmidt, J.P., Davies, C.J., Gollub, J., Webster, T., Wong, B., Zhan, Y., Auton, A., Campbell, C.L., Kong, Y., Marcketta, A., Gibbs, R.A., Yu, F., Antunes, L., Bainbridge, M., Muzny, D., Sabo, A., Huang, Z., Wang, J., Coin, L.J.M., Fang, L., Guo, X., Jin, X., Li, G., Li, Q., Li, Y., Li, Z., Lin, H., Liu, B., Luo, R., Shao, H., Xie, Y., Ye, C., Yu, C., Zhang, F., Zheng, H., Zhu, H., Alkan, C., Dal, E., Kahveci, F., Marth, G.T., Garrison, E.P., Kural, D., Lee, W.-P., Fung Leong, W., Stromberg, M.,

Ward, A.N., Wu, J., Zhang, M., Daly, M.J., DePristo, M.A., Handsaker, R.E., Altshuler, D.M., Banks, E., Bhatia, G., Angel, G. del, Gabriel, S.B., Genovese, G., Gupta, N., Li, H., Kashin, S., Lander, E.S., McCarroll, S.A., Nemesh, J.C., Poplin, R.E., Yoon, S.C., Lihm, J., Makarov, V., Clark, A.G., Gottipati, S., Keinan, A., Rodriguez-Flores, J.L., Korbel, J.O., Rausch, T., Fritz, M.H., Stütz, A.M., Flicek, P., Beal, K., Clarke, L., Datta, A., Herrero, J., McLaren, W.M., Ritchie, G.R.S., Smith, R.E., Zerbino, D., Zheng-Bradley, X., Sabeti, P.C., Shlyakhter, I., Schaffner, S.F., Vitti, J., Cooper, D.N., Ball, E.V., Stenson, P.D., Bentley, D.R., Barnes, B., Bauer, M., Keira Cheetham, R., Cox, A., Eberle, M., Humphray, S., Kahn, S., Murray, L., Peden, J., Shaw, R., Kenny, E.E., Batzer, M.A., Konkel, M.K., Walker, J.A., MacArthur, D.G., Lek, M., Sudbrak, R., Amstislavskiy, V.S., Herwig, R., Mardis, E.R., Ding, L., Koboldt, D.C., Larson, D., Ye, K., Gravel, S., Swaroop, A., Chew, E., Lappalainen, T., Erlich, Y., Gymrek, M., Frederick Willems, T., Simpson, J.T., Shriver, M.D., Rosenfeld, J.A., Bustamante, C.D., Montgomery, S.B., De La Vega, F.M., Byrnes, J.K., Carroll, A.W., DeGorter, M.K., Lacroute, P., Maples, B.K., Martin, A.R., Moreno-Estrada, A., Shringarpure, S.S., Zakharia, F., Halperin, E., Baran, Y., Lee, C., Cerveira, E., Hwang, J., Malhotra, A., Plewczynski, D., Radew, K., Romanovitch, M., Zhang, C., Hyland, F.C.L., Craig, D.W., Christoforides, A., Homer, N., Izatt, T., Kurdoglu, A.A., Sinari, S.A., Squire, K., Sherry, S.T., Xiao, C., Sebat, J., Antaki, D., Gujral, M., Noor, A., Ye, K., Burchard, E.G., Hernandez, R.D., Gignoux, C.R., Haussler, D., Katzman, S.J., James Kent, W., Howie, B., Ruiz-Linares, A., Dermitzakis, E.T., Devine, S.E., Abecasis, G.R., Min Kang, H., Kidd, J.M., Blackwell, T., Caron, S., Chen, W., Emery, S., Fritsche, L., Fuchsberger, C., Jun, G., Li, B., Lyons, R., Scheller, C., Sidore, C., Song, S., Sliwerska, E., Taliun, D., Tan, A., Welch, R., Kate Wing, M., Zhan, X., Awadalla, P., Hodgkinson, A., Li, Y., Shi, X., Quitadamo, A., Lunter, G., McVean, G.A., Marchini, J.L., Myers, S., Churchhouse, C., Delaneau, O., Gupta-Hinch, A., Kretzschmar, W., Iqbal, Z., Mathieson, I., Menelaou, A., Rimmer, A., Xifara, D.K., Oleksyk, T.K., Fu, Y., Liu, X., Xiong, M., Jorde, L., Witherspoon, D., Xing, J., Eichler, E.E., Browning, B.L., Browning, S.R., Hormozdiari, F., Sudmant, P.H., Khurana, E., Durbin, R.M., Hurles, M.E., Tyler-Smith, C., Albers, C.A., Ayub, Q., Balasubramaniam, S., Chen, Y., Colonna, V., Danecek, P., Jostins, L., Keane, T.M., McCarthy, S., Walter, K., Xue, Y., Gerstein, M.B., Abyzov, A., Balasubramanian, S., Chen, J., Clarke, D., Fu, Y., Harmanci, A.O., Jin, M., Lee, D., Liu, J., Jasmine Mu, X., Zhang, J., Zhang, Y., Li, Y., Luo, R., Zhu, H., Alkan, C., Dal, E., Kahveci, F., Marth, G.T., Garrison, E.P., Kural, D., Lee, W.-P., Ward, A.N., Wu, J., Zhang, M., McCarroll, S.A., Handsaker, R.E., Altshuler, D.M., Banks, E., Angel, G. del, Genovese, G., Hartl, C., Li, H., Kashin,

S., Nemesh, J.C., Shakir, K., Yoon, S.C., Lihm, J., Makarov, V., Degenhardt, J., Korbel, J.O., Fritz, M.H., Meiers, S., Raeder, B., Rausch, T., Stütz, A.M., Flicek, P., Paolo Casale, F., Clarke, L., Smith, R.E., Stegle, O., Zheng-Bradley, X., Bentley, D.R., Barnes, B., Keira Cheetham, R., Eberle, M., Humphray, S., Kahn, S., Murray, L., Shaw, R., Lameijer, E.-W., Batzer, M.A., Konkel, M.K., Walker, J.A., Ding, L., Hall, I., Ye, K., Lacroute, P., Lee, C., Cerveira, E., Malhotra, A., Hwang, J., Plewczynski, D., Radew, K., Romanovitch, M., Zhang, C., Craig, D.W., Homer, N., Church, D., Xiao, C., Sebat, J., Antaki, D., Bafna, V., Michaelson, J., Ye, K., Devine, S.E., Gardner, E.J., Abecasis, G.R., Kidd, J.M., Mills, R.E., Dayama, G., Emery, S., Jun, G., Shi, X., Quitadamo, A., Lunter, G., McVean, G.A., Chen, K., Fan, X., Chong, Z., Chen, T., Witherspoon, D., Xing, J., Eichler, E.E., Chaisson, M.J., Hormozdiari, F., Huddleston, J., Malig, M., Nelson, B.J., Sudmant, P.H., Parrish, N.F., Khurana, E., Hurles, M.E., Blackburne, B., Lindsay, S.J., Ning, Z., Walter, K., Zhang, Y., Gerstein, M.B., Abyzov, A., Chen, J., Clarke, D., Lam, H., Jasmine Mu, X., Sisu, C., Zhang, J., Zhang, Y., Gibbs, R.A., Yu, F., Bainbridge, M., Challis, D., Evani, U.S., Kovar, C., Lu, J., Muzny, D., Nagaswamy, U., Reid, J.G., Sabo, A., Yu, J., Guo, X., Li, W., Li, Y., Wu, R., Marth, G.T., Garrison, E.P., Fung Leong, W., Ward, A.N., Angel, G. del, DePristo, M.A., Gabriel, S.B., Gupta, N., Hartl, C., Poplin, R.E., Clark, A.G., Rodriguez-Flores, J.L., Flicek, P., Clarke, L., Smith, R.E., Zheng-Bradley, X., MacArthur, D.G., Mardis, E.R., Fulton, R., Koboldt, D.C., Gravel, S., Bustamante, C.D., Craig, D.W., Christoforides, A., Homer, N., Izatt, T., Sherry, S.T., Xiao, C., Dermitzakis, E.T., Abecasis, G.R., Min Kang, H., McVean, G.A., Gerstein, M.B., Balasubramanian, S., Habegger, L., Yu, H., Flicek, P., Clarke, L., Cunningham, F., Dunham, I., Zerbino, D., Zheng-Bradley, X., Lage, K., Berg Jespersen, J., Horn, H., Montgomery, S.B., DeGorter, M.K., Khurana, E., Tyler-Smith, C., Chen, Y., Colonna, V., Xue, Y., Gerstein, M.B., Balasubramanian, S., Fu, Y., Kim, D., Auton, A., Marcketta, A., Desalle, R., Narechania, A., Wilson Sayres, M.A., Garrison, E.P., Handsaker, R.E., Kashin, S., McCarroll, S.A., Rodriguez-Flores, J.L., Flicek, P., Clarke, L., Zheng-Bradley, X., Erlich, Y., Gymrek, M., Frederick Willems, T., Bustamante, C.D., Mendez, F.L., David Poznik, G., Underhill, P.A., Lee, C., Cerveira, E., Malhotra, A., Romanovitch, M., Zhang, C., Abecasis, G.R., Coin, L., Shao, H., Mittelman, D., Tyler-Smith, C., Ayub, Q., Banerjee, R., Cerezo, M., Chen, Y., Fitzgerald, T.W., Louzada, S., Massaia, A., McCarthy, S., Ritchie, G.R., Xue, Y., Yang, F., Gibbs, R.A., Kovar, C., Kalra, D., Hale, W., Muzny, D., Reid, J.G., Wang, J., Dan, X., Guo, X., Li, G., Li, Y., Ye, C., Zheng, X., Altshuler, D.M., Flicek, P., Clarke, L., Zheng-Bradley, X., Bentley, D.R., Cox, A., Humphray, S., Kahn, S., Sudbrak, R., Albrecht, M.W., Lienhard, M., Larson, D., Craig, D.W.,

Izatt, T., Kurdoglu, A.A., Sherry, S.T., Xiao, C., Haussler, D., Abecasis, G.R., McVean, G.A., Durbin, R.M., Balasubramaniam, S., Keane, T.M., McCarthy, S., Stalker, J., Chakravarti, A., Knoppers, B.M., Abecasis, G.R., Barnes, K.C., Beiswanger, C., Burchard, E.G., Bustamante, C.D., Cai, H., Cao, H., Durbin, R.M., Gerry, N.P., Gharani, N., Gibbs, R.A., Gignoux, C.R., Gravel, S., Henn, B., Jones, D., Jorde, L., Kaye, J.S., Keinan, A., Kent, A., Kerasidou, A., Li, Y., Mathias, R., McVean, G.A., Moreno-Estrada, A., Ossorio, P.N., Parker, M., Resch, A.M., Rotimi, C.N., Royal, C.D., Sandoval, K., Su, Y., Sudbrak, R., Tian, Z., Tishkoff, S., Toji, L.H., Tyler-Smith, C., Via, M., Wang, Y., Yang, H., Yang, L., Zhu, J., Bodmer, W., Bedoya, G., Ruiz-Linares, A., Cai, Z., Gao, Y., Chu, J., Peltonen, L., Garcia-Montero, A., Orfao, A., Dutil, J., Martinez-Cruzado, J.C., Oleksyk, T.K., Barnes, K.C., Mathias, R.A., Hennis, A., Watson, H., McKenzie, C., Qadri, F., LaRocque, R., Sabeti, P.C., Zhu, J., Deng, X., Sabeti, P.C., Asogun, D., Folarin, O., Happi, C., Omoniwa, O., Stremlau, M., Tariyal, R., Jallow, M., Sisay Joof, F., Corrah, T., Rockett, K., Kwiatkowski, D., Kooner, J., Tịnh Hiê‘n, T., Dunstan, S.J., Thuy Hang, N., Fonnie, R., Garry, R., Kanneh, L., Moses, L., Sabeti, P.C., Schieffelin, J., Grant, D.S., Gallo, C., Poletti, G., Saleheen, D., Rasheed, A., Brooks, L.D., Felsenfeld, A.L., McEwen, J.E., Vaydylevich, Y., Green, E.D., Duncanson, A., Dunn, M., Schloss, J.A., Wang, J., Yang, H., Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Min Kang, H., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., Abecasis, G.R., 2015. A global reference for human genetic variation. Nature 526, 68–74. https://doi.org/10.1038/nature15393

Babapour, B., Doustkami, H., Avesta, L., Moradi, A., Saadat, S., Piralaei, K., Aslani, M.R., 2020. Correlation of Serum Adipolin with Epicardial Fat Thickness and Severity of Coronary Artery Diseases in Acute Myocardial Infarction and Stable Angina Pectoris Patients. Medical Principles and Practice. https://doi.org/10.1159/000508834

Bagabir, R., Byers, R.J., Chaudhry, I.H., Müller, W., Paus, R., Bayat, A., 2012a. Site-specific immunophenotyping of keloid disease demonstrates immune upregulation and the presence of lymphoid aggregates. British Journal of Dermatology 167, 1053–1066. https://doi.org/10.1111/j.1365-2133.2012.11190.x

Bagabir, R., Syed, F., Paus, R., Bayat, A., 2012b. Long-term organ culture of keloid disease tissue. Experimental Dermatology 21, 376–381. https://doi.org/10.1111/j.1600-0625.2012.01476.x

Baharlou, H., Canete, N.P., Cunningham, A.L., Harman, A.N., Patrick, E., 2019. Mass cytometry imaging for the study of human diseases—applications and data analysis strategies. Frontiers in Immunology 10. https://doi.org/10.3389/

fimmu.2019.02657

Bai, B., Ban, B., Liu, Z., Zhang, M.M., Tan, B.K., Chen, J., 2017. Circulating C1q complement/TNF-related protein (CTRP) 1, CTRP9, CTRP12 and CTRP13 concentrations in Type 2 diabetes mellitus: In vivo regulation by glucose. PLOS ONE 12, e0172271. https://doi.org/10.1371/journal.pone.0172271

Baker, J., Riley, G., Romero, M.R., Haynes, A.R., Hilton, H., Simon, M., Hancock, J., Tateossian, H., Ripoll, V.M., Blanco, G., 2010. Identification of a Z-band associated protein complex involving KY, FLNC and IGFN1. Experimental Cell Research 316, 1856–1870. https://doi.org/10.1016/j.yexcr.2010.02.027

Barallobre-Barreiro, J., Woods, E., Bell, R.E., Easton, J.A., Hobbs, C., Eager, M., Baig, F., Ross, A.M., Mallipeddi, R., Powell, B., Soldin, M., Mayr, M., Shaw, T.J., 2019. Cartilage-like composition of keloid scar extracellular matrix suggests fibroblast mis-differentiation in disease. Matrix Biology Plus 4, 100016. https://doi.org/10.1016/j.mbplus.2019.100016

Baranyi, U., Winter, B., Gugerell, A., Hegedus, B., Brostjan, C., Laufer, G., Messner, B., 2019. Primary Human Fibroblasts in Culture Switch to a Myofibroblast-Like Phenotype Independently of TGF Beta. Cells 8, 721. https://doi.org/10.3390/cells8070721

Barros, M.H.M., Hauck, F., Dreyer, J.H., Kempkes, B., Niedobitek, G., 2013. Macrophage Polarisation: an Immunohistochemical Approach for Identifying M1 and M2 Macrophages. PLoS ONE 8, e80908. https://doi.org/10.1371/journal.pone.0080908

Bastarache, L., 2021. Using Phecodes for Research with the Electronic Health Record: From PheWAS to PheRS. Annual Review of Biomedical Data Science 4, 1–19. https://doi.org/10.1146/annurev-biodatasci-122320-112352

Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E.H., Britto, R., Bursteinas, B., Bye-A-Jee, H., Coetzee, R., Cukura, A., Da Silva, A., Denny, P., Dogan, T., Ebenezer, T., Fan, J., Castro, L.G., Garmiri, P., Georghiou, G., Gonzales, L., Hatton-Ellis, E., Hussein, A., Ignatchenko, A., Insana, G., Ishtiaq, R., Jokinen, P., Joshi, V., Jyothi, D., Lock, A., Lopez, R., Luciani, A., Luo, J., Lussi, Y., MacDougall, A., Madeira, F., Mahmoudy, M., Menchi, M., Mishra, A., Moulang, K., Nightingale, A., Oliveira, C.S., Pundir, S., Qi, G., Raj, S., Rice, D., Lopez, M.R., Saidi, R., Sampson, J., Sawford, T., Speretta, E., Turner, E., Tyagi, N., Vasudev, P., Volynkin, V., Warner, K., Watkins, X., Zaru, R., Zellner, H., Bridge, A., Poux, S., Redaschi, N., Aimo, L., Argoud-Puy, G., Auchincloss, A., Axelsen, K., Bansal, P., Baratin, D., Blatter, M.-C., Bolleman, J., Boutet, E., Breuza, L.,

Casals-Casas, C., Castro, E. de, Echioukh, K.C., Coudert, E., Cuche, B., Doche, M., Dornevil, D., Estreicher, A., Famiglietti, M.L., Feuermann, M., Gasteiger, E., Gehant, S., Gerritsen, V., Gos, A., Gruaz-Gumowski, N., Hinz, U., Hulo, C., Hyka-Nouspikel, N., Jungo, F., Keller, G., Kerhornou, A., Lara, V., Le Mercier, P., Lieberherr, D., Lombardot, T., Martin, X., Masson, P., Morgat, A., Neto, T.B., Paesano, S., Pedruzzi, I., Pilbout, S., Pourcel, L., Pozzato, M., Pruess, M., Rivoire, C., Sigrist, C., Sonesson, K., Stutz, A., Sundaram, S., Tognolli, M., Verbregue, L., Wu, C.H., Arighi, C.N., Arminski, L., Chen, C., Chen, Y., Garavelli, J.S., Huang, H., Laiho, K., McGarvey, P., Natale, D.A., Ross, K., Vinayaka, C.R., Wang, Q., Wang, Y., Yeh, L.-S., Zhang, J., Ruch, P., Teodoro, D., 2020. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research 49, D480–D489. https://doi.org/10.1093/nar/gkaa1100

Baum, C.L., Arpey, C.J., 2006. Normal Cutaneous Wound Healing: Clinical Correlation with Cellular and Molecular Events. Dermatologic Surgery 31, 674–686. https://doi.org/10.1111/j.1524-4725.2005.31612

Bayat, A., Arscott, G., Ollier, W.E.R., Mc Grouther, D.A., Ferguson, M.W.J., 2005a. Keloid disease: clinical relevance of single versus multiple site scars. British Journal of Plastic Surgery 58, 28–37. https://doi.org/10.1016/j.bjps.2004.04.024

Bayat, A., Bock, O., Mrowietz, U., Ollier, W.E.R., Ferguson, M.W.J., 2004. Genetic susceptibility to keloid disease: Transforming growth factor beta receptor gene polymorphisms are not associated with keloid disease. Experimental Dermatology 13, 120–124. https://doi.org/10.1111/j.0906-6705.2004.00165.x

Bayat, A., Bock, O., Mrowietz, U., Ollier, W.E.R., Ferguson, M.W.J., 2003. Genetic Susceptibility to Keloid Disease and Hypertrophic Scarring: Transforming Growth Factor ?1 Common Polymorphisms and Plasma Levels. Plastic and Reconstructive Surgery 111, 535–543. https://doi.org/10.1097/01.prs.0000041536.02524.a3

Bayat, A., Ollier, W.E.R., Ferguson, M.W.J., Bock, O., Mrowiet, U., 2002. Genetic susceptibility to keloid disease and transforming growth factor ?2 polymorphisms. British Journal of Plastic Surgery 55, 283–286. https://doi.org/10.1054/bjps.2002.3853

Bayat, A., Walter, J.M., Bock, O., Mrowietz, U., Ollier, W.E.R., Ferguson, M.W.J., 2005b. Genetic susceptibility to keloid disease: mutation screening of the TGF?3 gene. British Journal of Plastic Surgery 58, 914–921. https://doi.org/10.1016/j.bjps.2005.04.009

Bell, R.E., Shaw, T.J., 2021. Keloid tissue analysis discredits a role for myofi-

broblasts in disease pathogenesis. Wound Repair and Regeneration 29, 637–641. https://doi.org/10.1111/wrr.12923

Bell-Anderson, K.S., Funnell, A.P., Williams, H., Mat Jusoh, H., Scully, T., Lim, W.F., Burdach, J.G., Mak, K.S., Knights, A.J., Hoy, A.J., Nicholas, H.R., Sainsbury, A., Turner, N., Pearson, R.C., Crossley, M., 2013. Loss of Krüppel-Like Factor 3 (KLF3/BKLF) Leads to Upregulation of the Insulin-Sensitizing Factor Adipolin (FAM132A/CTRP12/C1qdc2). Diabetes 62, 2728–2737. https://doi.org/10.2337/db12-1745

Bettinger, D.A., Yager, D.R., Diegelmann, R.F., Cohen, K.I., 1996. The Effect of TGF-beta1 on Keloid Fibroblast Proliferation and Collagen Synthesis. Plastic and Reconstructive Surgery 98, 827–833.

Blair, David R., Lyttle, Christopher S., Mortensen, Jonathan M., Bearden, Charles F., Jensen, A., Khiabanian, H., Melamed, R., Rabadan, R., Bernstam, Elmer V., Brunak, S., Jensen, L., Nicolae, D., Shah, Nigam H., Grossman, Robert L., Cox, Nancy J., White, Kevin P., Rzhetsky, A., 2013. A Nondegenerate Code of Deleterious Variants in Mendelian Loci Contributes to Complex Disease Risk. Cell 155, 70–80. https://doi.org/10.1016/j.cell.2013.08.030

Blake, J.A., Baldarelli, R., Kadin, J.A., Richardson, J.E., Smith, C., Bult, C.J., Anagnostopoulos, A.V., Beal, J.S., Bello, S.M., Blodgett, O., Butler, N.E., Campbell, J., Christie, K.R., Corbani, L.E., Dolan, M.E., Drabkin, H.J., Flores, M., Giannatto, S.L., Guerra, A., Hale, P., Hill, D.P., Judd, J., Law, M., McAndrews, M., Miers, D., Mitchell, C., Motenko, H., Ni, L., Onda, H., Ormsby, J., Perry, M., Recla, J.M., Shaw, D., Sitnikov, D., Tomczuk, M., Wilming, L., Zhu, Y.'Sophia'., 2020. Mouse Genome Database (MGD): Knowledgebase for mouse–human comparative biology. Nucleic Acids Research 49, D981–D987. https://doi.org/10.1093/nar/gkaa1083

Bloom, D., 1956. Heredity of keloids; review of the literature and report of a family with multiple keloids in five generations. New York State Journal of Medicine 56, 511–519.

Blume-Peytavi, U., Geilen, C.C., Sommer, C., Almond-Roesler, B., Orfanos, C.E., 1997. The phospholipid analogue hexadecylphosphocholine (HePC) inhibits proliferation of keloid fibroblasts in vitro and modulates their fibronectin and integrin synthesis. Archives of Dermatological Research 289, 164–169. https://doi.org/10.1007/s004030050173

Boase, N.A., Kumar, S., 2015. NEDD4: The founding member of a family of ubiquitin-protein ligases. Gene 557, 113–122. https://doi.org/10.1016/j.gene.2014.12.020

Bobik, A., 2006. Transforming Growth Factor-?s and Vascular Disorders. Arteriosclerosis, Thrombosis, and Vascular Biology 26, 1712–1720. https://doi.org/10.1161/01.atv.0000225287.20034.2c

Bock, O., Yu, H., Zitron, S., Bayat, A., Ferguson, M., Mrowietz, U., 2005. Studies of transforming growth factors beta 1-3 and their receptors i and II in fibroblast of keloids and hypertrophic scars. Acta Dermato-Venereologica -1, 1–1. https://doi.org/10.1080/00015550410025453

Boland, M.R., Hripcsak, G., Albers, D.J., Wei, Y., Wilcox, A.B., Wei, J., Li, J., Lin, S., Breene, M., Myers, R., Zimmerman, J., Papapanou, P.N., Weng, C., 2013. Discovering medical conditions associated with periodontitis using linked electronic health records. Journal of Clinical Periodontology 40, 474–482. https://doi.org/10.1111/jcpe.12086

Borges, M.C., Barros, A.J.D., Ferreira, D.L.S., Casas, J.P., Horta, B.L., Kivimaki, M., Kumari, M., Menon, U., Gaunt, T.R., Ben-Shlomo, Y., Freitas, D.F., Oliveira, I.O., Gentry-Maharaj, A., Fourkala, E., Lawlor, D.A., Hingorani, A.D., 2017. Metabolic Profiling of Adiponectin Levels in Adults. Circulation: Cardiovascular Genetics 10. https://doi.org/10.1161/circgenetics.117.001837

Boyce, D.E., Ciampolini, J., Ruge, F., Harding, K.G., Murison, M.M.S.C., 2001. Inflammatory cell subpopulations in keloid scars. British Journal of Plastic Surgery 54, 511–516. https://doi.org/10.1054/bjps.2001.3638

Bozkurt, B., Aguilar, D., Deswal, A., Dunbar, S.B., Francis, G.S., Horwich, T., Jessup, M., Kosiborod, M., Pritchett, A.M., Ramasubbu, K., Rosendorff, C., Yancy, C., 2016. Contributory Risk and Management of Comorbidities of Hypertension, Obesity, Diabetes Mellitus, Hyperlipidemia, and Metabolic Syndrome in Chronic Heart Failure: A Scientific Statement From the American Heart Association. Circulation 134. https://doi.org/10.1161/cir.0000000000000450

Bradford, J.R., Cox, A., Bernard, P., Camp, N.J., 2016. Consensus Analysis of Whole Transcriptome Profiles from Two Breast Cancer Patient Cohorts Reveals Long Non-Coding RNAs Associated with Intrinsic Subtype and the Tumour Microenvironment. PLOS ONE 11, e0163238. https://doi.org/10.1371/journal.pone.0163238

Brown, J.J., Ollier, W., Arscott, G., Ke, X., Lamb, J., Day, P., Bayat, A., 2008a. Genetic susceptibility to Keloid scarring: SMAD gene SNP frequencies in Afro-Caribbeans. Experimental Dermatology 17, 610–613. https://doi.org/10.1111/j.1600-0625.2007.00654.x

Brown, J.J., Ollier, W.E.R., Thomson, W., Bayat, A., 2008b. Positive association of HLA-DRB1*15 with keloid disease in Caucasians. International Jour-

nal of Immunogenetics 35, 303–307. https://doi.org/10.1111/j.1744-313x.2008.00780.x

Bujor, A.M., Pannu, J., Bu, S., Smith, E.A., Muise-Helmericks, R.C., Trojanowska, M., 2008. Akt Blockade Downregulates Collagen and Upregulates MMP1 in Human Dermal Fibroblasts. Journal of Investigative Dermatology 128, 1906–1914. https://doi.org/10.1038/jid.2008.39

Bulik-Sullivan, B.K., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.-R., Duncan, L., Perry, J.R.B., Patterson, N., Robinson, E.B., Daly, M.J., Price, A.L., Neale, B.M., 2015a. An atlas of genetic correlations across human diseases and traits. Nature Genetics 47, 1236–1241. https://doi.org/10.1038/ng.3406

Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Patterson, N., Daly, M.J., Price, A.L., Neale, B.M., 2015b. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nature Genetics 47, 291–295. https://doi.org/10.1038/ng.3211

Burd, A., Huang, L., 2005. Hypertrophic Response and Keloid Diathesis: Two Very Different Forms of Scar. Plastic and Reconstructive Surgery 116, 150e–157e. https://doi.org/10.1097/01.prs.0000191977.51206.43

Burk, C.J., Aber, C., Connelly, E.A., 2007. Ehlers-Danlos syndrome type IV: Keloidal plaques of the lower extremities, amniotic band limb deformity, and a new mutation. Journal of the American Academy of Dermatology 56, S53–S54. https://doi.org/10.1016/j.jaad.2006.11.008

Butzelaar, L., Schooneman, D.P.M., Soykan, E.A., Talhout, W., Ulrich, M.M.W., Broek, L.J. van den, Gibbs, S., Beelen, R.H.J., Mink van der Molen, A.B., Niessen, F.B., 2016. Inhibited early immunologic response is associated with hypertrophic scarring. Experimental Dermatology 25, 797–804. https://doi.org/10.1111/exd.13100

Buuren, S. van, Groothuis-Oudshoorn, K., 2011. {Mice}: Multivariate imputation by chained equations in r 45, 1–67. https://doi.org/10.18637/jss.v045.i03

Bux, S., Madaree, A., 2010. Keloids Show Regional Distribution of Proliferative and Degenerate Connective Tissue Elements. Cells Tissues Organs 191, 213–234. https://doi.org/10.1159/000231899

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P., Marchini, J., 2018. The UK Biobank resource with deep phenotyping and genomic data. Nature 562, 203–209. https://doi.org/10.1038/s41586-018-0579-z

Calippe, B., Douin-Echinard, V., Delpy, L., Laffargue, M., Lélu, K., Krust, A.,

Pipy, B., Bayard, F., Arnal, J.-F., Guéry, J.-C., Gourdy, P., 2010. 17 -Estradiol Promotes TLR4-Triggered Proinflammatory Mediator Production through Direct Estrogen Receptor Signaling in Macrophages In Vivo. The Journal of Immunology 185, 1169–1176. https://doi.org/10.4049/jimmunol.0902383

Cano-Gamez, E., Trynka, G., 2020. From GWAS to function: Using functional genomics to identify the mechanisms underlying complex diseases. Frontiers in Genetics 11. https://doi.org/10.3389/fgene.2020.00424

Carroll, R.J., Bastarache, L., Denny, J.C., 2014. R {PheWAS}: Data analysis and plotting tools for phenome-wide association studies in the r environment 30, 2375–2376. https://doi.org/10.1093/bioinformatics/btu197

Carter, A.B., Abruzzo, L.V., Hirschhorn, J.W., Jones, D., Jordan, D.C., Nassiri, M., Ogino, S., Patel, N.R., Suciu, C.G., Temple-Smolkin, R.L., Zehir, A., Roy, S., 2022. Electronic Health Records and Genomics. The Journal of Molecular Diagnostics 24, 1–17. https://doi.org/10.1016/j.jmoldx.2021.09.009

Casanova, J.-L., Abel, L., Quintana-Murci, L., 2011. Human TLRs and IL-1Rs in Host Defense: Natural Insights from Evolutionary, Epidemiological, and Clinical Genetics. Annual Review of Immunology 29, 447–491. https://doi.org/10.1146/annurev-immunol-030409-101335

Censin, J.C., Bovijn, J., Holmes, M.V., Lindgren, C.M., 2021. Colocalization analysis of polycystic ovary syndrome to identify potential disease-mediating genes and proteins. European Journal of Human Genetics 29, 1446–1454. https://doi.org/10.1038/s41431-021-00835-8

Chaudet, K.M., Goyal, A., Veprauskas, K.R., Nazarian, R.M., 2020. Wnt Signaling Pathway Proteins in Scar, Hypertrophic Scar, and Keloid: Evidence for a Continuum? The American Journal of Dermatopathology 42, 842–847. https://doi.org/10.1097/dad.0000000000001661

Chen, C.S., Mrksich, M., Huang, S., Whitesides, G.M., Ingber, D.E., 1997. Geometric Control of Cell Life and Death. Science 276, 1425–1428. https://doi.org/10.1126/science.276.5317.1425

Chen, W.-H., Zhao, X.-M., Noort, V. van, Bork, P., 2013. Human Monogenic Disease Genes Have Frequently Functionally Redundant Paralogs. PLoS Computational Biology 9, e1003073. https://doi.org/10.1371/journal.pcbi.1003073

Chen, Y., Gao, J.-H., Liu, X.-J., Yan, X., Song, M., 2006a. Characteristics of occurrence for Han Chinese familial keloids. Burns 32, 1052–1059. https://doi.org/10.1016/j.burns.2006.04.014

Chen, Y., Gao, J., Liu, X., Yan, X., Song, M., 2006b. [Linkage analysis of keloid susceptibility loci on chromosome 7p11 in a Chinese pedigree]. Nan Fang Yi Ke

Da Xue Xue Bao = Journal of Southern Medical University 26, 623–625, 637.

Chen, Y., Ohki, R., 2020. p53-PHLDA3-Akt Network: The Key Regulators of Neuroendocrine Tumorigenesis. International Journal of Molecular Sciences 21, 4098. https://doi.org/10.3390/ijms21114098

Chipev, C.C., Simman, R., Hatch, G., Katz, A.E., Siegel, D.M., Simon, M., 2000. Myofibroblast phenotype and apoptosis in keloid and palmar fibroblasts in vitro. Cell Death & Differentiation 7, 166–176. https://doi.org/10.1038/sj.cdd.4400605

Chipev, C.C., Simon, M., 2002. Phenotypic differences between dermal fibroblasts from different body sites determine their responses to tension and TGF?1. BMC Dermatology 2. https://doi.org/10.1186/1471-5945-2-13

Clark, J.A., Turner, M.L., Howard, L., Stanescu, H., Kleta, R., Kopp, J.B., 2009. Description of familial keloids in five pedigrees: evidence for autosomal dominant inheritance and phenotypic heterogeneity. BMC Dermatology 9. https://doi.org/10.1186/1471-5945-9-8

Clayton, S.W., Ban, G.I., Liu, C., Serra, R., 2020. Canonical and noncanonical TGF-signaling regulate fibrous tissue differentiation in the axial skeleton. Scientific Reports 10. https://doi.org/10.1038/s41598-020-78206-4

Conde, J., Scotece, M., López, V., Gómez, R., Lago, F., Pino, J., Gómez-Reino, J.J., Gualillo, O., 2012. Expression and modulation of adipolin/C1qdc2: a novel adipokine in human and murine ATDC-5 chondrocyte cell line. Annals of the Rheumatic Diseases 72, 140–142.

Connor, J.W., Gomez, E.W., 2013. Cell Adhesion and Shape Regulate TGF-Beta1-Induced Epithelial-Myofibroblast Transition via MRTF-A Signaling. PLoS ONE 8, e83188. https://doi.org/10.1371/journal.pone.0083188

Cunningham, F., Allen, J.E., Allen, J., Alvarez-Jarreta, J., Amode, M., Armean, I., Austine-Orimoloye, O., Azov, A., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., Davidson, C., Dodiya, K., Donaldson, S., El Houdaigui, B., El Naboulsi, T., Fatima, R., Giron, C.G., Genez, T., Martinez, J., Guijarro-Clarke, C., Gymer, A., Hardy, M., Hollis, Z., Hourlier, T., Hunt, T., Juettemann, T., Kaikala, V., Kay, M., Lavidas, I., Le, T., Lemos, D., Marugán, J.C., Mohanan, S., Mushtaq, A., Naven, M., Ogeh, D., Parker, A., Parton, A., Perry, M., Piližota, I., Prosovetskaia, I., Sakthivel, M., Salam, A., Schmitt, B., Schuilenburg, H., Sheppard, D., Pérez-Silva, J., Stark, W., Steed, E., Sutinen, K., Sukumaran, R., Sumathipala, D., Suner, M.-M., Szpak, M., Thormann, A., Tricomi, F.F., Urbina-Gómez, D., Veidenberg, A., Walsh, T., Walts, B.,

Willhoft, N., Winterbottom, A., Wass, E., Chakiachvili, M., Flint, B., Frankish, A., Giorgetti, S., Haggerty, L., Hunt, S., IIsley, G., Loveland, J., Martin, F., Moore, B., Mudge, J., Muffato, M., Perry, E., Ruffier, M., Tate, J., Thybert, D., Trevanion, S., Dyer, S., Harrison, P., Howe, K., Yates, A., Zerbino, D., Flicek, P., 2021. Ensembl 2022. Nucleic Acids Research 50, D988–D995. https://doi.org/10.1093/nar/gkab1049

Daian, T., Ishihara, H., Hirano, A., Fujii, T., Ohtsuru, A., Rogounovitch, T., Akiyama-Uchida, Y., Saenko, V., Yamashita, S., 2003. Insulin-Like Growth Factor-I Enhances Transforming Growth Factor-?-Induced Extracellular Matrix Protein Production Through the P38/Activating Transcription Factor-2 Signaling Pathway in Keloid Fibroblasts. Journal of Investigative Dermatology 120, 956–962. https://doi.org/10.1046/j.1523-1747.2003.12143.x

Darmawan, C.C., Montenegro, S.E., Jo, G., Kusumaningrum, N., Lee, S.-H., Chung, J.-H., Mun, J.-H., 2020. Adiponectin-Based Peptide (ADP355) Inhibits Transforming Growth Factor- 1-Induced Fibrosis in Keloids. International Journal of Molecular Sciences 21, 2833. https://doi.org/10.3390/ijms21082833

DeBoever, C., Tanigawa, Y., Lindholm, M.E., McInnes, G., Lavertu, A., Ingelsson, E., Chang, C., Ashley, E.A., Bustamante, C.D., Daly, M.J., Rivas, M.A., 2018. Medical relevance of protein-truncating variants across 337,205 individuals in the UK Biobank study. Nature Communications 9. https://doi.org/10.1038/s41467-018-03910-9

Deng, C.-C., Hu, Y.-F., Zhu, D.-H., Cheng, Q., Gu, J.-J., Feng, Q.-L., Zhang, L.-X., Xu, Y.-P., Wang, D., Rong, Z., Yang, B., 2021. Single-cell RNA-seq reveals fibroblast heterogeneity and increased mesenchymal fibroblasts in human fibrotic skin diseases. Nature Communications 12. https://doi.org/10.1038/s41467-021-24110-y

Deng, Y., Scherer, P.E., 2010. Adipokines as novel biomarkers and regulators of the metabolic syndrome. Annals of the New York Academy of Sciences 1212, E1–E19. https://doi.org/10.1111/j.1749-6632.2010.05875.x

Denny, J.C., 2012. Chapter 13: Mining Electronic Health Records in the Genomics Era. PLoS Computational Biology 8, e1002823. https://doi.org/10.1371/journal.pcbi.1002823

Denny, J.C., Bastarache, L., Ritchie, M.D., Carroll, R.J., Zink, R., Mosley, J.D., Field, J.R., Pulley, J.M., Ramirez, A.H., Bowton, E., Basford, M.A., Carrell, D.S., Peissig, P.L., Kho, A.N., Pacheco, J.A., Rasmussen, L.V., Crosslin, D.R., Crane, P.K., Pathak, J., Bielinski, S.J., Pendergrass, S.A., Xu, H., Hindorff, L.A., Li, R., Manolio, T.A., Chute, C.G., Chisholm, R.L., Larson, E.B., Jarvik, G.P.,

Brilliant, M.H., McCarty, C.A., Kullo, I.J., Haines, J.L., Crawford, D.C., Masys, D.R., Roden, D.M., 2013. Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. Nature Biotechnology 31, 1102–1111. https://doi.org/10.1038/nbt.2749

desJardins-Park, H.E., Chinta, M.S., Foster, D.S., Borrelli, M.R., Shen, A.H., Wan, D.C., Longaker, M.T., 2020. Fibroblast Heterogeneity in and Its Implications for Plastic and Reconstructive Surgery. Plastic and Reconstructive Surgery - Global Open 8, e2927. https://doi.org/10.1097/gox.0000000000002927

Diaz, A., Tan, K., He, H., Xu, H., Cueto, I., Pavel, A.B., Krueger, J.G., Guttman-Yassky, E., 2020. Keloid lesions show increased IL -4/ IL -13 signaling and respond to Th2-targeting dupilumab therapy. Journal of the European Academy of Dermatology and Venereology 34. https://doi.org/10.1111/jdv.16097

Dmytrzak, A., Boroń, A., Łoniewska, B., Lewandowska, K., Gorący, I., Kaczmarczyk, M., Ciechanowicz, A., 2019. Two Functional *TP53* Genetic Variants and Predisposition to Keloid Scarring in Caucasians. Dermatology Research and Practice 2019, 1–5. https://doi.org/10.1155/2019/6179063

Dolan, M.E., Hill, D.P., Mukherjee, G., McAndrews, M.S., Chesler, E.J., Blake, J.A., 2020. Investigation of COVID-19 comorbidities reveals genes and pathways coincident with the SARS-CoV-2 viral disease. Scientific Reports 10. https://doi.org/10.1038/s41598-020-77632-8

Driskell, R.R., Lichtenberger, B.M., Hoste, E., Kretzschmar, K., Simons, B.D., Charalambous, M., Ferron, S.R., Herault, Y., Pavlovic, G., Ferguson-Smith, A.C., Watt, F.M., 2013. Distinct fibroblast lineages determine dermal architecture in skin development and repair. Nature 504, 277–281. https://doi.org/10.1038/nature12783

Driskell, R.R., Watt, F.M., 2015. Understanding fibroblast heterogeneity in the skin. Trends in Cell Biology 25, 92–99. https://doi.org/10.1016/j.tcb.2014.10.001

Durinck, S., Spellman, P.T., Birney, E., Huber, W., 2009. Mapping identifiers for the integration of genomic datasets with the r/bioconductor package biomaRt 4.

Dustan, H.P., 1995. Does Keloid Pathogenesis Hold the Key to Understanding Black/White Differences in Hypertension Severity? Hypertension 26, 858–862. https://doi.org/10.1161/01.hyp.26.6.858

Edgar, R., 2002. Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Research 30, 207–210. https://doi.org/10.1093/nar/30.1.207

Edqvist, P.-H.D., Fagerberg, L., Hallström, B.M., Danielsson, A., Edlund, K., Uhlén,

M., Pontén, F., 2014. Expression of Human Skin-Specific Genes Defined by Transcriptomics and Antibody-Based Profiling. Journal of Histochemistry & Cytochemistry 63, 129–141. https://doi.org/10.1369/0022155414562646

El Hadidi, H.H., Sobhi, R.M., Nada, A.M., AbdelGhaffar, M.M.M., Shaker, O.G., El-Kalioby, M., 2021. Does vitamin D deficiency predispose to keloids via dysregulation of koebnerisin (S100A15)? A case-control study. Wound Repair and Regeneration 29, 425–431. https://doi.org/10.1111/wrr.12894

Eming, S.A., Krieg, T., Davidson, J.M., 2007. Inflammation in Wound Repair: Molecular and Cellular Mechanisms. Journal of Investigative Dermatology 127, 514–525. https://doi.org/10.1038/sj.jid.5700701

Enomoto, T., Ohashi, K., Shibata, R., Higuchi, A., Maruyama, S., Izumiya, Y., Walsh, K., Murohara, T., Ouchi, N., 2011. Adipolin/C1qdc2/CTRP12 Protein Functions as an Adipokine That Improves Glucose Metabolism. Journal of Biological Chemistry 286, 34552–34558. https://doi.org/10.1074/jbc.m111.277319

Enomoto, T., Ohashi, K., Shibata, R., Kambara, T., Uemura, Y., Yuasa, D., Kataoka, Y., Miyabe, M., Matsuo, K., Joki, Y., Hayakawa, S., Hiramatsu-Ito, M., Ito, M., Murohara, T., Ouchi, N., 2013. Transcriptional Regulation of an Insulin-Sensitizing Adipokine Adipolin/CTRP12 in Adipocytes by Krüppel-Like Factor 15. PLoS ONE 8, e83183. https://doi.org/10.1371/journal.pone.0083183

Enomoto, T., Shibata, R., Ohashi, K., Kambara, T., Kataoka, Y., Uemura, Y., Yuasa, D., Murohara, T., Ouchi, N., 2012. Regulation of adipolin/CTRP12 cleavage by obesity. Biochemical and Biophysical Research Communications 428, 155–159. https://doi.org/10.1016/j.bbrc.2012.10.031

Evanko, S.P., Potter-Perigo, S., Petty, L.J., Workman, G.A., Wight, T.N., 2015. Hyaluronan Controls the Deposition of Fibronectin and Collagen and Modulates TGF-?1 Induction of Lung Myofibroblasts. Matrix Biology 42, 74–92. https://doi.org/10.1016/j.matbio.2014.12.001

Fadaei, R., Moradi, N., Kazemi, T., Chamani, E., Azdaki, N., Moezibady, S.A., Shahmohamadnejad, S., Fallah, S., 2019. Decreased serum levels of CTRP12/adipolin in patients with coronary artery disease in relation to inflammatory cytokines and insulin resistance. Cytokine 113, 326–331. https://doi.org/10.1016/j.cyto.2018.09.019

Fang, F., Huang, R.-L., Zheng, Y., Liu, M., Huo, R., 2016. Bone marrow derived mesenchymal stem cells inhibit the proliferative and profibrotic phenotype of hypertrophic scar fibroblasts and keloid fibroblasts through paracrine signaling.

Journal of Dermatological Science 83, 95–105. https://doi.org/10.1016/j.jdermsci.2016.03.003

Fang, F., Liu, L., Yang, Y., Tamaki, Z., Wei, J., Marangoni, R.G., Bhattacharyya, S., Summer, R.S., Ye, B., Varga, J., 2012. The adipokine adiponectin has potent anti-fibrotic effects mediated via adenosine monophosphate-activated protein kinase: novel target for fibrosis therapy. Arthritis Research & Therapy 14, R229. https://doi.org/10.1186/ar4070

Faresjö, T., Faresjö, Å., 2010. To Match or Not to Match in Epidemiological Studies—Same Outcome but Less Power. International Journal of Environmental Research and Public Health 7, 325–332. https://doi.org/10.3390/ijerph7010325

Feiglin, A., Allen, B.K., Kohane, I.S., Kong, S.W., 2017. Comprehensive Analysis of Tissue-wide Gene Expression and Phenotype Data Reveals Tissues Affected in Rare Genetic Disorders. Cell Systems 5, 140–148.e2. https://doi.org/10.1016/j.cels.2017.06.016

Feng, J., Xue, S., Pang, Q., Rang, Z., Cui, F., 2017. miR-141-3p inhibits fibroblast proliferation and migration by targeting GAB1 in keloids. Biochemical and Biophysical Research Communications 490, 302–308. https://doi.org/10.1016/j.bbrc.2017.06.040

Ferkingstad, E., Sulem, P., Atlason, B.A., Sveinbjornsson, G., Magnusson, M.I., Styrmisdottir, E.L., Gunnarsdottir, K., Helgason, A., Oddsson, A., Halldorsson, B.V., Jensson, B.O., Zink, F., Halldorsson, G.H., Masson, G., Arnadottir, G.A., Katrinardottir, H., Juliusson, K., Magnusson, M.K., Magnusson, O.Th., Fridriksdottir, R., Saevarsdottir, S., Gudjonsson, S.A., Stacey, S.N., Rognvaldsson, S., Eiriksdottir, T., Olafsdottir, T.A., Steinthorsdottir, V., Tragante, V., Ulfarsson, M.O., Stefansson, H., Jonsdottir, I., Holm, H., Rafnar, T., Melsted, P., Saemundsdottir, J., Norddahl, G.L., Lund, S.H., Gudbjartsson, D.F., Thorsteinsdottir, U., Stefansson, K., 2021. Large-scale integration of the plasma proteome with genetics and disease. Nature Genetics 53, 1712–1721. https://doi.org/10.1038/s41588-021-00978-w

Ferreira, M.A., Vonk, J.M., Baurecht, H., Marenholz, I., Tian, C., Hoffman, J.D., Helmer, Q., Tillander, A., Ullemar, V., Dongen, J. van, Lu, Y., Rüschendorf, F., Esparza-Gordillo, J., Medway, C.W., Mountjoy, E., Burrows, K., Hummel, O., Grosche, S., Brumpton, B.M., Witte, J.S., Hottenga, J.-J., Willemsen, G., Zheng, J., Rodríguez, E., Hotze, M., Franke, A., Revez, J.A., Beesley, J., Matheson, M.C., Dharmage, S.C., Bain, L.M., Fritsche, L.G., Gabrielsen, M.E., Balliu, B., Nielsen, J.B., Zhou, W., Hveem, K., Langhammer, A., Holmen, O.L., Løset, M., Abecasis, G.R., Willer, C.J., Arnold, A., Homuth, G., Schmidt, C.O., Thompson,

P.J., Martin, N.G., Duffy, D.L., Novak, N., Schulz, H., Karrasch, S., Gieger, C., Strauch, K., Melles, R.B., Hinds, D.A., Hübner, N., Weidinger, S., Magnusson, P.K.E., Jansen, R., Jorgenson, E., Lee, Y.-A., Boomsma, D.I., Almqvist, C., Karlsson, R., Koppelman, G.H., Paternoster, L., 2017. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. Nature Genetics 49, 1752–1757. https://doi.org/10.1038/ng.3985

Firth, D., 1993. Bias reduction of maximum likelihood estimates. Biometrika 80, 27–38. https://doi.org/10.1093/biomet/80.1.27

Fisher, G.J., Shao, Y., He, T., Qin, Z., Perry, D., Voorhees, J.J., Quan, T., 2015. Reduction of fibroblast size/mechanical force down-regulates TGF - type II receptor: implications for human skin aging. Aging Cell 15, 67–76. https://doi.org/10.1111/acel.12410

Fisher, H.M., Hoehndorf, R., Bazelato, B.S., Dadras, S.S., King, L.E., Gkoutos, G.V., Sundberg, J.P., Schofield, P.N., 2016. DermO; an ontology for the description of dermatologic disease. Journal of Biomedical Semantics 7. https://doi.org/10.1186/s13326-016-0085-x

Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., Lancet, D., Cohen, D., 2017. GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database 2017. https://doi.org/10.1093/database/bax028

Fitzgerald O'Connor, E.J., Badshah, I.I., Addae, L.Y., Kundasamy, P., Thanabalasingam, S., Abioye, D., Soldin, M., Shaw, T.J., 2012. Histone Deacetylase 2 Is Upregulated in Normal and Keloid Scars. Journal of Investigative Dermatology 132, 1293–1296. https://doi.org/10.1038/jid.2011.432

Fong, E.P., Bay, B.H., 2002. Keloids – the sebum hypothesis revisited. Medical Hypotheses 58, 264–269. https://doi.org/10.1054/mehy.2001.1426

Fong, E.P., Chye, L.T., Tan, W.T.L., 1999. Keloids: Time to dispel the myths? Plastic and Reconstructive Surgery 104, 1199–1202. https://doi.org/10.1097/00006534-199909020-00065

Freeberg, M.A., Fromont, L.A., D'Altri, T., Romero, A.F., Ciges, J., Jene, A., Kerry, G., Moldes, M., Ariosa, R., Bahena, S., Barrowdale, D., Barbero, M., Fernandez-Orth, D., Garcia-Linares, C., Garcia-Rios, E., Haziza, F., Juhasz, B., Llobet, O., Milla, G., Mohan, A., Rueda, M., Sankar, A., Shaju, D., Shimpi, A., Singh, B., Thomas, C., de la Torre, S., Uyan, U., Vasallo, C., Flicek, P., Guigo, R., Navarro, A., Parkinson, H., Keane, T., Rambla, J., 2021. The European Genome-phenome Archive in 2021. Nucleic Acids Research 50, D980–D987. https://doi.org/10.1093/nar/gkab1059

French, J.D., Edwards, S.L., 2020. The Role of Noncoding Variants in Heritable Disease. Trends in Genetics 36, 880–891. https://doi.org/10.1016/j.tig.2020.07.004

Froese, A.R., Shimbori, C., Bellaye, P.-S., Inman, M., Obex, S., Fatima, S., Jenkins, G., Gauldie, J., Ask, K., Kolb, M., 2016. Stretch-induced Activation of Transforming Growth Factor-?1in Pulmonary Fibrosis. American Journal of Respiratory and Critical Care Medicine 194, 84–96. https://doi.org/10.1164/rccm.201508-1638oc

Fromme, M., Schneider, C.V., Schlapbach, C., Cazzaniga, S., Trautwein, C., Rader, D.J., Borradori, L., Strnad, P., 2022. Comorbidities in lichen planus by phenome-wide association study in two biobank population cohorts. British Journal of Dermatology. https://doi.org/10.1111/bjd.21762

Fry, A., Littlejohns, T.J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., Allen, N.E., 2017. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. American Journal of Epidemiology 186, 1026–1034. https://doi.org/10.1093/aje/kwx246

Fujiwara, M., Muragaki, Y., Ooshima, A., 2005. Upregulation of transforming growth factor-?1 and vascular endothelial growth factor in cultured keloid fibroblasts: relevance to angiogenic activity. Archives of Dermatological Research 297, 161–169. https://doi.org/10.1007/s00403-005-0596-2

Fukushima, T., Yoshihara, H., Furuta, H., Kamei, H., Hakuno, F., Luan, J., Duan, C., Saeki, Y., Tanaka, K., Iemura, S.-I., Natsume, T., Chida, K., Nakatsu, Y., Kamata, H., Asano, T., Takahashi, S.-I., 2015. Nedd4-induced monoubiquitination of IRS-2 enhances IGF signalling and mitogenic activity. Nature Communications 6. https://doi.org/10.1038/ncomms7780

Gao, F., Cai, Y., Kapranov, P., Xu, D., 2020. Reverse-genetics studies of lncRNAs—what we have learnt and paths forward. Genome Biology 21. https://doi.org/10.1186/s13059-020-01994-5

Gauglitz, G.G., Korting, H.C., Pavicic, T., Ruzicka, T., Jeschke, M.G., 2010. Hypertrophic Scarring and Keloids: Pathomechanisms and Current and Emerging Treatment Strategies. Molecular Medicine 17, 113–125. https://doi.org/10.2119/molmed.2009.00153

Gavrila, A., Chan, J.L., Yiannakouris, N., Kontogianni, M., Miller, L.C., Orlova, C., Mantzoros, C.S., 2003. Serum Adiponectin Levels Are Inversely Associated with Overall and Central Fat Distribution but Are Not Directly Regulated by Acute Fasting or Leptin Administration in Humans: Cross-Sectional and Interventional

Studies. The Journal of Clinical Endocrinology & Metabolism 88, 4823–4831. https://doi.org/10.1210/jc.2003-030214

Geschickter, C.F., Lewis, D., 1935. Tumors of connective tissue. The American Journal of Cancer 25, 630–655. https://doi.org/10.1158/ajc.1935.630

Ghazizadeh, M., Tosa, M., Shimizu, H., Hyakusoku, H., Kawanami, O., 2007. Functional Implications of the IL-6 Signaling Pathway in Keloid Pathogenesis. Journal of Investigative Dermatology 127, 98–105. https://doi.org/10.1038/sj.jid.5700564

Giambartolomei, C., al, et., 2014. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics.

Giral, H., Landmesser, U., Kratzer, A., 2018. Into the wild: GWAS exploration of non-coding RNAs. Frontiers in Cardiovascular Medicine 5. https://doi.org/10.3389/fcvm.2018.00181

Giugliano, G., Pasquali, D., Notaro, A., Brongo, S., Nicoletti, G., D'Andrea, F., Bellastella, A., Sinisi, A.A., 2003. Verapamil inhibits interleukin-6 and vascular endothelial growth factor production in primary cultures of keloid fibroblasts. British Journal of Plastic Surgery 56, 804–809. https://doi.org/10.1016/s0007-1226(03)00384-9

Glass, D.A., 2017. Current Understanding of the Genetic Causes of Keloid Formation. Journal of Investigative Dermatology Symposium Proceedings 18, S50–S53. https://doi.org/10.1016/j.jisp.2016.10.024

Glim, J.E., Niessen, F.B., Everts, V., Egmond, M. van, Beelen, R.H.J., 2013. Platelet derived growth factor-CC secreted by M2 macrophages induces alpha-smooth muscle actin expression by dermal and gingival fibroblasts. Immunobiology 218, 924–929. https://doi.org/10.1016/j.imbio.2012.10.004

Goeminne, L., 1968. A new probably X-linked inherited syndrome: congenital muscular torticollis, multiple keloids cryptorchidism and renal dysplasia. Acta Geneticae Medicae Et Gemellologiae 17, 439–467. https://doi.org/10.1017/s1120962300012634

Gohel, D., 2022. Flextable: Functions for tabular reporting.

Gonza´lez-Marti´nez, R., Mari´n-Bertoli´n, S., Amorrortu-Velayos, J., 1995. Association between keloids and Dupuytren's disease: case report. British Journal of Plastic Surgery 48, 47–48. https://doi.org/10.1016/0007-1226(95)90031-4

Goodfellow, A., Emmerson, R.W., Calvert, H.T., 1980. Rubinstein-Taybi syndrome and spontaneous keloids. Clinical and Experimental Dermatology 5, 369–370. https://doi.org/10.1111/j.1365-2230.1980.tb01718.x

Goudarzi, M., Berg, K., Pieper, L.M., Schier, A.F., 2019. Individual long non-coding

RNAs have no overt functions in zebrafish embryogenesis, viability and fertility. eLife 8. https://doi.org/10.7554/elife.40815

Goumans, M.-J., 2002. Balancing the activation state of the endothelium via two distinct TGF-beta type i receptors. The EMBO Journal 21, 1743–1753. https://doi.org/10.1093/emboj/21.7.1743

Grabham, P., Sharma, P., 2013. The effects of radiation on angiogenesis. Vascular Cell 5, 19. https://doi.org/10.1186/2045-824x-5-19

Gregoire, F.M., 2001. Adipocyte Differentiation: From Fibroblast to Endocrine Cell. Experimental Biology and Medicine 226, 997–1002. https://doi.org/10.1177/153537020122601106

Griffin, M.F., desJardins-Park, H.E., Mascharak, S., Borrelli, M.R., Longaker, M.T., 2020. Understanding the impact of fibroblast heterogeneity on skin fibrosis. Disease Models & Mechanisms 13. https://doi.org/10.1242/dmm.044164

Grunewald, J., Brynedal, B., Darlington, P., Nisell, M., Cederlund, K., Hillert, J., Eklund, A., 2010. Different HLA-DRB1 allele distributions in distinct clinical subgroups of sarcoidosis patients. Respiratory Research 11. https://doi.org/10.1186/1465-9921-11-25

GTEx Consortium, 2017. Genetic effects on gene expression across human tissues. Nature 550, 204–213. https://doi.org/10.1038/nature24277

Gulamhuseinwala, N., Mackey, S., Meagher, P., Powell, B., 2008. Should Excised Keloid Scars Be Sent for Routine Histologic Analysis? Annals of Plastic Surgery 60, 186–187. https://doi.org/10.1097/sap.0b013e318056d6cc

Guo, H., Fortune, M.D., Burren, O.S., Schofield, E., Todd, J.A., Wallace, C., 2015. Integration of disease association and eQTL data using a Bayesian colocalisation approach highlights six candidate causal genes in immune-mediated diseases. Human Molecular Genetics 24, 3305–3313. https://doi.org/10.1093/hmg/ddv077

Guo, L., Xu, K., Yan, H., Feng, H., Chai, L., Xu, G., 2016. Expression Profile of Long Noncoding RNAs in Human Earlobe Keloids: A Microarray Analysis. BioMed Research International 2016, 1–9. https://doi.org/10.1155/2016/5893481

Haisa, M., Okochi, H., Grotendorst, G.R., 1994. Elevated Levels of PDGF ? Receptors in Keloid Fibroblasts Contribute to an Enhanced Response to PDGF. Journal of Investigative Dermatology 103, 560–563. https://doi.org/10.1111/1523-1747.ep12396856

Hallock, G.G., Rice, D.C., Merkel, J.R., DiPaolo, B.R., 1988. Analysis of Collagen Content in the Fetal Wound. Annals of Plastic Surgery 21, 310–315. https:

//doi.org/10.1097/00000637-198810000-00003

Han, X., Luo, S., Peng, G., Lu, J.Y., Cui, G., Liu, L., Yan, P., Yin, Y., Liu, W., Wang, R., Zhang, J., Ai, S., Chang, Z., Na, J., He, A., Jing, N., Shen, X., 2018. Mouse knockout models reveal largely dispensable but context-dependent functions of lncRNAs during development. Journal of Molecular Cell Biology 10, 175–178. https://doi.org/10.1093/jmcb/mjy003

Hanasono, M.M., Kita, M., Mikulec, A.A., Lonergan, D., Koch, R.J., 2003. Autocrine Growth Factor Production by Fetal, Keloid, and Normal Dermal Fibroblasts. Archives of Facial Plastic Surgery 5, 26–30. https://doi.org/10.1001/archfaci.5.1.26

Harmon, Q.E., Laughlin, S.K., Baird, D.D., 2013. Keloids and Ultrasound Detected Fibroids in Young African American Women. PLoS ONE 8, e84737. https://doi.org/10.1371/journal.pone.0084737

Hasegawa, T., Nakao, A., Sumiyoshi, K., Tsuboi, R., Ogawa, H., 2003. IFN- fails to antagonize fibrotic effect of TGF- on keloid-derived dermal fibroblasts. Journal of Dermatological Science 32, 19–24. https://doi.org/10.1016/s0923-1811(03)00044-6

Hawash, A.A., Ingrasci, G., Nouri, K., Yosipovitch, G., 2021. Pruritus in keloid scars: Mechanisms and treatments. Acta Dermato-Venereologica 101, adv00582. https://doi.org/10.2340/00015555-3923

He, F., Jacobson, A., 2015. Nonsense-Mediated mRNA Decay: Degradation of Defective Transcripts Is Only Part of the Story. Annual Review of Genetics 49, 339–366. https://doi.org/10.1146/annurev-genet-112414-054639

He, S., Liu, X., Yang, Y., Huang, W., Xu, S., Yang, S., Zhang, X., Roberts, M.S., 2009. Mechanisms of transforming growth factor $_1$ /Smad signalling mediated by mitogen-activated protein kinase pathways in keloid fibroblasts. British Journal of Dermatology 162, 538–546. https://doi.org/10.1111/j.1365-2133.2009.09511.x

Hebbring, S.J., 2014. The challenges, advantages and future of phenome-wide association studies. Immunology 141, 157–165. https://doi.org/10.1111/imm.12195

Hellstrom, L., Wahrenberg, H., Hruska, K., Reynisdottir, S., Arner, P., 2000. Mechanisms behind gender differences in circulating leptin levels. Journal of Internal Medicine 247, 457–462. https://doi.org/10.1046/j.1365-2796.2000.00678.x

Hesketh, M., Sahin, K.B., West, Z.E., Murray, R.Z., 2017. Macrophage Phenotypes Regulate Scar Formation and Chronic Wound Healing. International Journal of

Molecular Sciences 18, 1545. https://doi.org/10.3390/ijms18071545

Hill, W.G., Robertson, A., 1968. Linkage disequilibrium in finite populations. Theoretical and Applied Genetics 38, 226–231. https://doi.org/10.1007/bf01245622

Hinrichs, A.S., 2006. The UCSC Genome Browser Database: update 2006. Nucleic Acids Research 34, D590–D598. https://doi.org/10.1093/nar/gkj144

Hoerst, K., van den Broek, L., Sachse, C., Klein, O., von Fritschen, U., Gibbs, S., Hedtrich, S., 2019. Regenerative potential of adipocytes in hypertrophic scars is mediated by myofibroblast reprogramming. Journal of Molecular Medicine (Berlin, Germany) 97, 761–775. https://doi.org/10.1007/s00109-019-01772-2

Holick, M.F., Binkley, N.C., Bischoff-Ferrari, H.A., Gordon, C.M., Hanley, D.A., Heaney, R.P., Murad, M.H., Weaver, C.M., 2011. Evaluation, Treatment, and Prevention of Vitamin D Deficiency: an Endocrine Society Clinical Practice Guideline. The Journal of Clinical Endocrinology & Metabolism 96, 1911–1930. https://doi.org/10.1210/jc.2011-0385

Hu, J.X., Thomas, C.E., Brunak, S., 2016. Network biology concepts in complex disease comorbidities. Nature Reviews Genetics 17, 615–629. https://doi.org/10.1038/nrg.2016.87

Hu, P., Jiao, R., Jin, L., Xiong, M., 2018. Application of causal inference to genomic analysis: Advances in methodology. Frontiers in Genetics 9. https://doi.org/10.3389/fgene.2018.00238

Hu, Z.-C., Shi, F., Liu, P., Zhang, J., Guo, D., Cao, X.-L., Chen, C.-F., Qu, S.-Q., Zhu, J.-Y., Tang, B., 2017. TIEG1 Represses Smad7-Mediated Activation of TGF- 1/Smad Signaling in Keloid Pathogenesis. Journal of Investigative Dermatology 137, 1051–1059. https://doi.org/10.1016/j.jid.2016.12.019

Huang, C., Akaishi, S., Hyakusoku, H., Ogawa, R., 2012. Are keloid and hypertrophic scar different forms of the same disorder? A fibroproliferative skin disorder hypothesis based on keloid findings. International Wound Journal 11, 517–522. https://doi.org/10.1111/j.1742-481x.2012.01118.x

Huang, C., Ogawa, R., 2013. Roles of lipid metabolism in keloid development. Lipids in Health and Disease 12. https://doi.org/10.1186/1476-511x-12-60

Huang, F., Wang, K., Shen, J., 2019. Lipoprotein-associated phospholipase A2: The story continues. Medicinal Research Reviews 40, 79–134. https://doi.org/10.1002/med.21597

Hutchinson, A., Wallace, C., 2019. Corrcoverage: Correcting the coverage of credible sets from bayesian genetic fine mapping.

Ibrahim, N.E., Shaharan, S., Dheansa, B., 2020. Adverse Effects of Pregnancy on

Keloids and Hypertrophic Scars. Cureus. https://doi.org/10.7759/cureus.12154

Igota, S., Tosa, M., Murakami, M., Egawa, S., Shimizu, H., Hyakusoku, H., Ghazizadeh, M., 2013. Identification and Characterization of Wnt Signaling Pathway in Keloid Pathogenesis. International Journal of Medical Sciences 10, 344–354. https://doi.org/10.7150/ijms.5349

Im, M.J., Hoopes, J.E., 1970. Energy metabolism in healing skin wounds. The Journal of Surgical Research 10, 459–464. https://doi.org/10.1016/0022-4804(70)90070-3

Ishigaki, K., Akiyama, M., Kanai, M., Takahashi, A., Kawakami, E., Sugishita, H., Sakaue, S., Matoba, N., Low, S.-K., Okada, Y., Terao, C., Amariuta, T., Gazal, S., Kochi, Y., Horikoshi, M., Suzuki, K., Ito, K., Koyama, S., Ozaki, K., Niida, S., Sakata, Y., Sakata, Y., Kohno, T., Shiraishi, K., Momozawa, Y., Hirata, M., Matsuda, K., Ikeda, M., Iwata, N., Ikegawa, S., Kou, I., Tanaka, T., Nakagawa, H., Suzuki, A., Hirota, T., Tamari, M., Chayama, K., Miki, D., Mori, M., Nagayama, S., Daigo, Y., Miki, Y., Katagiri, T., Ogawa, O., Obara, W., Ito, H., Yoshida, T., Imoto, I., Takahashi, T., Tanikawa, C., Suzuki, T., Sinozaki, N., Minami, S., Yamaguchi, H., Asai, S., Takahashi, Y., Yamaji, K., Takahashi, K., Fujioka, T., Takata, R., Yanai, H., Masumoto, A., Koretsune, Y., Kutsumi, H., Higashiyama, M., Murayama, S., Minegishi, N., Suzuki, K., Tanno, K., Shimizu, A., Yamaji, T., Iwasaki, M., Sawada, N., Uemura, H., Tanaka, K., Naito, M., Sasaki, M., Wakai, K., Tsugane, S., Yamamoto, M., Yamamoto, K., Murakami, Y., Nakamura, Y., Raychaudhuri, S., Inazawa, J., Yamauchi, T., Kadowaki, T., Kubo, M., Kamatani, Y., 2020. Large scale genome-wide association study in a japanese population identifies novel susceptibility loci across different diseases. Nature genetics 52, 669–679. https://doi.org/10.1038/s41588-020-0640-3

Jiang, D., Rinkevich, Y., 2020. Scars or Regeneration?—Dermal Fibroblasts as Drivers of Diverse Skin Wound Responses. International Journal of Molecular Sciences 21, 617. https://doi.org/10.3390/ijms21020617

Jiang, L., Zheng, Z., Qi, T., Kemper, K.E., Wray, N.R., Visscher, P.M., Yang, J., 2019. A resource-efficient tool for mixed model association analysis of large-scale data. Nature Genetics 51, 1749–1755.

Jin, J., Jia, Z.-H., Luo, X.-H., Zhai, H.-F., 2019. Long non-coding RNA HOXA11-AS accelerates the progression of keloid formation via miR-124-3p/TGF R1 axis. Cell Cycle 19, 218–232. https://doi.org/10.1080/15384101.2019.1706921

Jin, Q., Gui, L., Niu, F., Yu, B., Lauda, N., Liu, J., Mao, X., Chen, Y., 2018. Macrophages in keloid are potent at promoting the differentiation and function

of regulatory T cells. Experimental Cell Research 362, 472–476. https://doi.org/10.1016/j.yexcr.2017.12.011

Johnson, J.L., Abecasis, G.R., 2017. GAS power calculator: Web-based power calculator for genetic association studies. bioRxiv 164343. https://doi.org/10.1101/164343

Jones, L.R., Young, W., Divine, G., Datta, I., Chen, K.M., Ozog, D., Worsham, M.J., 2015. Genome-Wide Scan for Methylation Profiles in Keloids. Disease Markers 2015, 1–7. https://doi.org/10.1155/2015/943176

Jumper, N., Paus, R., Bayat, T., 2015. Functional histopathology of keloid disease. Histology and Histopathology 1033–1057. https://doi.org/10.14670/HH-11-624

Kamamoto, F., Paggiaro, A.O., Rodas, A., Herson, M.R., Mathor, M.B., Ferreira, M.C., 2003. A Wound Contraction Experimental Model for Studying Keloids and Wound-healing Modulators. Artificial Organs 27, 701–705. https://doi.org/10.1046/j.1525-1594.2003.07277.x

Kamat, M.A., Blackshaw, J.A., Young, R., Surendran, P., Burgess, S., Danesh, J., Butterworth, A.S., Staley, J.R., 2019. PhenoScanner V2: an expanded tool for searching human genotype–phenotype associations. Bioinformatics 35, 4851–4853. https://doi.org/10.1093/bioinformatics/btz469

Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S., Freimer, N.B., Sabatti, C., Eskin, E., 2010. Variance component model to account for sample structure in genome-wide association studies. Nature Genetics 42, 348–354. https://doi.org/10.1038/ng.548

Kao, J.S., Fluhr, J.W., Man, M.-Q., Fowler, A.J., Hachem, J.-P., Crumrine, D., Ahn, S.K., Brown, B.E., Elias, P.M., Feingold, K.R., 2003. Short-Term Glucocorticoid Treatment Compromises Both Permeability Barrier Homeostasis and Stratum Corneum Integrity: Inhibition of Epidermal Lipid Synthesis Accounts for Functional Abnormalities. Journal of Investigative Dermatology 120, 456–464. https://doi.org/10.1046/j.1523-1747.2003.12053.x

Karbowiak, M., Holme, T., Khan, K., Mohan, A., 2021. Dupuytren's disease. BMJ n1308. https://doi.org/10.1136/bmj.n1308

Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., Gauthier, L.D., Brand, H., Solomonson, M., Watts, N.A., Rhodes, D., Singer-Berk, M., England, E.M., Seaby, E.G., Kosmicki, J.A., Walters, R.K., Tashman, K., Farjoun, Y., Banks, E., Poterba, T., Wang, A., Seed, C., Whiffin, N., Chong, J.X., Samocha, K.E., Pierce-Hoffman, E., Zappala, Z., O'Donnell-Luria, A.H., Minikel, E.V.,

Weisburd, B., Lek, M., Ware, J.S., Vittal, C., Armean, I.M., Bergelson, L., Cibulskis, K., Connolly, K.M., Covarrubias, M., Donnelly, S., Ferriera, S., Gabriel, S., Gentry, J., Gupta, N., Jeandet, T., Kaplan, D., Llanwarne, C., Munshi, R., Novod, S., Petrillo, N., Roazen, D., Ruano-Rubio, V., Saltzman, A., Schleicher, M., Soto, J., Tibbetts, K., Tolonen, C., Wade, G., Talkowski, M.E., Aguilar Salinas, C.A., Ahmad, T., Albert, C.M., Ardissino, D., Atzmon, G., Barnard, J., Beaugerie, L., Benjamin, E.J., Boehnke, M., Bonnycastle, L.L., Bottinger, E.P., Bowden, D.W., Bown, M.J., Chambers, J.C., Chan, J.C., Chasman, D., Cho, J., Chung, M.K., Cohen, B., Correa, A., Dabelea, D., Daly, M.J., Darbar, D., Duggirala, R., Dupuis, J., Ellinor, P.T., Elosua, R., Erdmann, J., Esko, T., Färkkilä, M., Florez, J., Franke, A., Getz, G., Glaser, B., Glatt, S.J., Goldstein, D., Gonzalez, C., Groop, L., Haiman, C., Hanis, C., Harms, M., Hiltunen, M., Holi, M.M., Hultman, C.M., Kallela, M., Kaprio, J., Kathiresan, S., Kim, B.-J., Kim, Y.J., Kirov, G., Kooner, J., Koskinen, S., Krumholz, H.M., Kugathasan, S., Kwak, S.H., Laakso, M., Lehtimäki, T., Loos, R.J.F., Lubitz, S.A., Ma, R.C.W., MacArthur, D.G., Marrugat, J., Mattila, K.M., McCarroll, S., McCarthy, M.I., McGovern, D., McPherson, R., Meigs, J.B., Melander, O., Metspalu, A., Neale, B.M., Nilsson, P.M., O'Donovan, M.C., Ongur, D., Orozco, L., Owen, M.J., Palmer, C.N.A., Palotie, A., Park, K.S., Pato, C., Pulver, A.E., Rahman, N., Remes, A.M., Rioux, J.D., Ripatti, S., Roden, D.M., Saleheen, D., Salomaa, V., Samani, N.J., Scharf, J., Schunkert, H., Shoemaker, M.B., Sklar, P., Soininen, H., Sokol, H., Spector, T., Sullivan, P.F., Suvisaari, J., Tai, E.S., Teo, Y.Y., Tiinamaija, T., Tsuang, M., Turner, D., Tusie-Luna, T., Vartiainen, E., Vawter, M.P., Ware, J.S., Watkins, H., Weersma, R.K., Wessman, M., Wilson, J.G., Xavier, R.J., Neale, B.M., Daly, M.J., MacArthur, D.G., 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581, 434–443. https://doi.org/10.1038/s41586-020-2308-7

Karczewski, K.J., Solomonson, M., Chao, K.R., Goodrich, J.K., Tiao, G., Lu, W., Riley-Gillis, B.M., Tsai, E.A., Kim, H.I., Zheng, X., Rahimov, F., Esmaeeli, S., Grundstad, A.J., Reppell, M., Waring, J., Jacob, H., Sexton, D., Bronson, P.G., Chen, X., Hu, X., Goldstein, J.I., King, D., Vittal, C., Poterba, T., Palmer, D.S., Churchhouse, C., Howrigan, D.P., Zhou, W., Watts, N.A., Nguyen, K., Nguyen, H., Mason, C., Farnham, C., Tolonen, C., Gauthier, L.D., Gupta, N., MacArthur, D.G., Rehm, H.L., Seed, C., Philippakis, A.A., Daly, M.J., Davis, J.W., Runz, H., Miller, M.R., Neale, B.M., 2022. Systematic single-variant and gene-based association testing of thousands of phenotypes in 426,370 UK biobank exomes. medRxiv 2021.06.19.21259117. https://doi.org/10.1101/2021.06.19.21259117

Karczewski, K.J., Weisburd, B., Thomas, B., Solomonson, M., Ruderfer, D.M., Kavanagh, D., Hamamsy, T., Lek, M., Samocha, K.E., Cummings, B.B., Birnbaum, D., Daly, M.J., MacArthur, D.G., 2016. The ExAC browser: displaying reference data information from over 60 000 exomes. Nucleic Acids Research 45, D840–D845. https://doi.org/10.1093/nar/gkw971

Karlsson, M., Zhang, C., Méar, L., Zhong, W., Digre, A., Katona, B., Sjöstedt, E., Butler, L., Odeberg, J., Dusart, P., Edfors, F., Oksvold, P., Feilitzen, K. von, Zwahlen, M., Arif, M., Altay, O., Li, X., Ozcan, M., Mardinoglu, A., Fagerberg, L., Mulder, J., Luo, Y., Ponten, F., Uhlén, M., Lindskog, C., 2021. A single–cell type transcriptomics map of human tissues. Science Advances 7. https://doi.org/10.1126/sciadv.abh2169

Kasabri, V., Al-ghareeb, M.I., Saleh, M.I., Suyagh, M., Halaseh, L., Al-Sarraf, I., AlAlawi, S., 2019. Proportional correlates of adipolin and cathepsin S in metabolic syndrome patients with and without prediabetes. Diabetes & Metabolic Syndrome: Clinical Research & Reviews 13, 2403–2408. https://doi.org/10.1016/j.dsx.2019.06.010

Kaukonen, R., Jacquemet, G., Hamidi, H., Ivaska, J., 2017. Cell-derived matrices for studying cell proliferation and directional migration in a complex 3D microenvironment. Nature Protocols 12, 2376–2390. https://doi.org/10.1038/nprot.2017.107

Kawase, T., Ohki, R., Shibata, T., Tsutsumi, S., Kamimura, N., Inazawa, J., Ohta, T., Ichikawa, H., Aburatani, H., Tashiro, F., Taya, Y., 2009. PH Domain-Only Protein PHLDA3 Is a p53-Regulated Repressor of Akt. Cell 136, 535–550. https://doi.org/10.1016/j.cell.2008.12.002

Kervestin, S., Jacobson, A., 2012. NMD: a multifaceted response to premature translational termination. Nature Reviews Molecular Cell Biology 13, 700–712. https://doi.org/10.1038/nrm3454

Khoo, Y.T., Ong, C.T., Mukhopadhyay, A., Han, H.C., Do, D.V., Lim, I.J., Phan, T.T., 2006. Upregulation of secretory connective tissue growth factor (CTGF) in keratinocyte-fibroblast coculture contributes to keloid pathogenesis. Journal of Cellular Physiology 208, 336–343. https://doi.org/10.1002/jcp.20668

Kieran, I., Knock, A., Bush, J., So, K., Metcalfe, A., Hobson, R., Mason, T., O'Kane, S., Ferguson, M., 2013. Interleukin-10 reduces scar formation in both animal and human cutaneous wounds: Results of two preclinical and phase II randomized control studies. Wound Repair and Regeneration 21, 428–436. https://doi.org/10.1111/wrr.12043

Kilian, K.A., Bugarija, B., Lahn, B.T., Mrksich, M., 2010. Geometric cues for

directing the differentiation of mesenchymal stem cells. Proceedings of the National Academy of Sciences 107, 4872–4877. https://doi.org/10.1073/pnas.0903269107

Kiprono, S.K., Chaula, B.M., Masenga, J.E., Muchunu, J.W., Mavura, D.R., Moehrle, M., 2015. Epidemiology of keloids in normally pigmented Africans and African people with albinism: population-based cross-sectional survey. British Journal of Dermatology 173, 852–854. https://doi.org/10.1111/bjd.13826

Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., Shendure, J., 2014. A general framework for estimating the relative pathogenicity of human genetic variants. Nature Genetics 46, 310–315. https://doi.org/10.1038/ng.2892

Klingberg, F., Chow, M.L., Koehler, A., Boo, S., Buscemi, L., Quinn, T.M., Costell, M., Alman, B.A., Genot, E., Hinz, B., 2014. Prestress in the extracellular matrix sensitizes latent TGF-?1 for activation. Journal of Cell Biology 207, 283–297. https://doi.org/10.1083/jcb.201402006

Koike, S., Akaishi, S., Nagashima, Y., Dohi, T., Hyakusoku, H., Ogawa, R., 2014. Nd. Plastic and Reconstructive Surgery Global Open 2, e272. https://doi.org/10.1097/gox.0000000000000231

Kolde, R., 2019. Pheatmap: Pretty heatmaps.

Kölling, M., Genschel, C., Kaucsar, T., Hübner, A., Rong, S., Schmitt, R., Sörensen-Zender, I., Haddad, G., Kistler, A., Seeger, H., Kielstein, J.T., Fliser, D., Haller, H., Wüthrich, R., Zörnig, M., Thum, T., Lorenzen, J., 2018. Hypoxia-induced long non-coding RNA Malat1 is dispensable for renal ischemia/reperfusion-injury. Scientific Reports 8. https://doi.org/10.1038/s41598-018-21720-3

Koni, M., Pinnarò, V., Brizzi, M.F., 2020. The Wnt Signalling Pathway: A Tailored Target in Cancer. International Journal of Molecular Sciences 21, 7697. https://doi.org/10.3390/ijms21207697

Kuan, V., Denaxas, S., Gonzalez-Izquierdo, A., Direk, K., Bhatti, O., Husain, S., Sutaria, S., Hingorani, M., Nitsch, D., Parisinos, C.A., Lumbers, R.T., Mathur, R., Sofat, R., Casas, J.P., Wong, I.C.K., Hemingway, H., Hingorani, A.D., 2019. A chronological map of 308 physical and mental health conditions from 4 million individuals in the English National Health Service. The Lancet Digital Health 1, e63–e77. https://doi.org/10.1016/s2589-7500(19)30012-3

Kulminski, A.M., Loika, Y., Culminskaya, I., Arbeev, K.G., Ukraintseva, S.V., Stallard, E., Yashin, A.I., 2016. Explicating heterogeneity of complex traits has strong potential for improving GWAS efficiency. Scientific Reports 6. https://doi.org/10.1038/srep35390

Kurita, M., Okazaki, M., Kaminishi-Tanikawa, A., Niikura, M., Takushima, A.,

Harii, K., 2012. Differential Expression of Wound Fibrotic Factors between Facial and Trunk Dermal Fibroblasts. Connective Tissue Research 53, 349–354. https://doi.org/10.3109/03008207.2012.657309

Kurki, M.I., Karjalainen, J., Palta, P., Sipilä, T.P., Kristiansson, K., Donner, K., Reeve, M.P., Laivuori, H., Aavikko, M., Kaunisto, M.A., Loukola, A., Lahtela, E., Mattsson, H., Laiho, P., Della Briotta Parolo, P., Lehisto, A., Kanai, M., Mars, N., Rämö, J., Kiiskinen, T., Heyne, H.O., Veerapen, K., Rüeger, S., Lemmelä, S., Zhou, W., Ruotsalainen, S., Pärn, K., Hiekkalinna, T., Koskelainen, S., Paajanen, T., Llorens, V., Gracia-Tabuenca, J., Siirtola, H., Reis, K., Elnahas, A.G., Aalto-Setälä, K., Alasoo, K., Arvas, M., Auro, K., Biswas, S., Bizaki-Vallaskangas, A., Carpen, O., Chen, C.-Y., Dada, O.A., Ding, Z., Ehm, M.G., Eklund, K., Färkkilä, M., Finucane, H., Ganna, A., Ghazal, A., Graham, R.R., Green, E., Hakanen, A., Hautalahti, M., Hedman, Å., Hiltunen, M., Hinttala, R., Hovatta, I., Hu, X., Huertas-Vazquez, A., Huilaja, L., Hunkapiller, J., Jacob, H., Jensen, J.-N., Joensuu, H., John, S., Julkunen, V., Jung, M., Junttila, J., Kaarniranta, K., Kähönen, M., Kajanne, R.M., Kallio, L., Kälviäinen, R., Kaprio, J., Kerimov, N., Kettunen, J., Kilpeläinen, E., Kilpi, T., Klinger, K., Kosma, V.-M., Kuopio, T., Kurra, V., Laisk, T., Laukkanen, J., Lawless, N., Liu, A., Longerich, S., Mägi, R., Mäkelä, J., Mäkitie, A., Malarstig, A., Mannermaa, A., Maranville, J., Matakidou, A., Meretoja, T., Mozaffari, S.V., Niemi, M.EK., Niemi, M., Niiranen, T., O'Donnell, C.J., Obeidat, M., Okafo, G., Ollila, H.M., Palomäki, A., Palotie, T., Partanen, J., Paul, D.S., Pelkonen, M., Pendergrass, R.K., Petrovski, S., Pitkäranta, A., Platt, A., Pulford, D., Punkka, E., Pussinen, P., Raghavan, N., Rahimov, F., Rajpal, D., Renaud, N.A., Riley-Gillis, B., Rodosthenous, R., Saarentaus, E., Salminen, A., Salminen, E., Salomaa, V., Schleutker, J., Serpi, R., Shen, H., Siegel, R., Silander, K., Siltanen, S., Soini, S., Soininen, H., Sul, J.H., Tachmazidou, I., Tasanen, K., Tienari, P., Toppila-Salmi, S., Tukiainen, T., Tuomi, T., Turunen, J.A., Ulirsch, J.C., Vaura, F., Virolainen, P., Waring, J., Waterworth, D., Yang, R., Nelis, M., Reigo, A., Metspalu, A., Milani, L., Esko, T., Fox, C., Havulinna, A.S., Perola, M., Ripatti, S., Jalanko, A., Laitinen, T., Mäkelä, T., Plenge, R., McCarthy, M., Runz, H., Daly, M.J., Palotie, A., 2022. FinnGen: Unique genetic insights from combining isolated population and national health register data. medRxiv 2022.03.03.22271360. https://doi.org/10.1101/2022.03.03.22271360

Kurokawa, N., Ueda, K., Tsuji, M., 2010. Study of microvascular structure in keloid and hypertrophic scars: Density of microvessels and the efficacy of three-dimensional vascular imaging. Journal of Plastic Surgery and Hand Surgery 44, 272–277. https://doi.org/10.3109/2000656x.2010.532923

Kurwa, A.R., 1979. Rubinstein-Taybi syndrome and spontaneous keloids. Clinical and Experimental Dermatology 4, 251–254. https://doi.org/10.1111/j.1365-2230.1979.tb01627.x

Kuwahara, H., Tosa, M., Egawa, S., Murakami, M., Mohammad, G., Ogawa, R., 2016. Examination of Epithelial Mesenchymal Transition in Keloid Tissues and Possibility of Keloid Therapy Target. Plastic and Reconstructive Surgery - Global Open 4, e1138. https://doi.org/10.1097/gox.0000000000001138

Kwon, H.-E., Ahn, H.-J., Jeong, S.J., Shin, M.K., 2021. The increased prevalence of keloids in atopic dermatitis patients with allergic comorbidities: a nationwide retrospective cohort study. Scientific Reports 11. https://doi.org/10.1038/s41598-021-03164-4

Lachmann, A., Torre, D., Keenan, A.B., Jagodnik, K.M., Lee, H.J., Wang, L., Silverstein, M.C., Ma'ayan, A., 2018. Massive mining of publicly available RNA-seq data from human and mouse. Nature Communications 9. https://doi.org/10.1038/s41467-018-03751-6

Lacombe, D., Morice-Picard, F., 2014. Rare genetic diseases, signalling pathways, and keloid scar formation. British Journal of Dermatology 171, 452–453. https://doi.org/10.1111/bjd.13257

Ladin, D.A., Hou, Z., Patel, D., McPhail, M., Olson, J.C., Saed, G.M., Fivenson, D.P., 1998. p53 and apoptosis alterations in keloids and keloid fibroblasts. Wound Repair and Regeneration 6, 28–37. https://doi.org/10.1046/j.1524-475x.1998.60106.x

Lah, M., Niranjan, T., Srikanth, S., Holloway, L., Schwartz, C.E., Wang, T., Weaver, D.D., 2016. A distinct X-linked syndrome involving joint contractures, keloids, large optic cup-to-disc ratio, and renal stones results from a filamin A (*FLNA*) mutation. American Journal of Medical Genetics Part A 170, 881–890. https://doi.org/10.1002/ajmg.a.37567

Lakota, K., Wei, J., Carns, M., Hinchcliff, M., Lee, J., Whitfield, M.L., Sodin-Semrl, S., Varga, J., 2012. Levels of adiponectin, a marker for PPAR-gamma activity, correlate with skin fibrosis in systemic sclerosis: potential utility as a biomarker? Arthritis Research & Therapy 14, R102. https://doi.org/10.1186/ar3827

Landau, W.M., 2021. The targets r package: A dynamic make-like function-oriented pipeline toolkit for reproducibility and high-performance computing 6, 2959.

Lappalainen, T., MacArthur, D.G., 2021. From variant to function in human disease genetics. Science 373, 1464–1468. https://doi.org/10.1126/science.abi8207

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R.,

Morgan, M., Carey, V., 2013. Software for computing and annotating genomic ranges 9. https://doi.org/10.1371/journal.pcbi.1003118

Lawton Snyder, A., Zmuda, JosephM., Thompson, PaulD., 1996. Keloid associated with hypertension. The Lancet 347, 465–466. https://doi.org/10.1016/s0140-6736(96)90042-2

Lebeko, M., Khumalo, N.P., Bayat, A., 2019. Multi-dimensional models for functional testing of keloid scars: In silico, in vitro, organoid, organotypic, ex vivo organ culture, and in vivo models. Wound Repair and Regeneration 27, 298–308. https://doi.org/10.1111/wrr.12705

Lee, J.Y.-Y., Yang, C.-C., Chao, S.-C., Wong, T.-W., 2004. Histopathological Differential Diagnosis of Keloid and Hypertrophic Scar. The American Journal of Dermatopathology 26, 379–384.

Lee, S., Abecasis, Gonçalo R., Boehnke, M., Lin, X., 2014. Rare-Variant Association Analysis: Study Designs and Statistical Tests. The American Journal of Human Genetics 95, 5–23. https://doi.org/10.1016/j.ajhg.2014.06.009

Lee, W.J., Choi, I.-K., Lee, J.H., Kim, Y.O., Yun, C.-O., 2012. A novel three-dimensional model system for keloid study: Organotypic multicellular scar model. Wound Repair and Regeneration 21, 155–165. https://doi.org/10.1111/j.1524-475x.2012.00869.x

Lee, Y.-S., Hsu, T., Chiu, W.-C., Sarkozy, H., Kulber, D.A., Choi, A., Kim, E.W., Benya, P.D., Tuan, T.-L., 2016. Keloid-derived, plasma/fibrin-based skin equivalents generate de novo dermal and epidermal pathology of keloid fibrosis in a mouse model. Wound Repair and Regeneration 24, 302–316. https://doi.org/10.1111/wrr.12397

Leeuw, C.A. de, Mooij, J.M., Heskes, T., Posthuma, D., 2015. MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLOS Computational Biology 11, e1004219. https://doi.org/10.1371/journal.pcbi.1004219

Li, B., Ritchie, M.D., 2021. From GWAS to gene: Transcriptome-wide association studies and other methods to functionally understand GWAS discoveries. Frontiers in Genetics 12. https://doi.org/10.3389/fgene.2021.713230

Li, G., Li, Y.-Y., Sun, J.-E., Lin, W., Zhou, R., 2016. ILK–PI3K/AKT pathway participates in cutaneous wound contraction by regulating fibroblast migration and differentiation to myofibroblast. Laboratory Investigation 96, 741–751. https://doi.org/10.1038/labinvest.2016.48

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/

bioinformatics/btp352

Li, J., Cao, J., Li, M., Yu, Y., Yang, Y., Xiao, X., Wu, Z., Wang, L., Tu, Y., Chen, H., 2011. Collagen triple helix repeat containing-1 inhibits transforming growth factor-?1-induced collagen type I expression in keloid. British Journal of Dermatology 164, 1030–1036. https://doi.org/10.1111/j.1365-2133.2011.10215.x

Li, X., Baker, J., Cracknell, T., Haynes, A.R., Blanco, G., 2017. IGFN1_v1 is required for myoblast fusion and differentiation. PLOS ONE 12, e0180217. https://doi.org/10.1371/journal.pone.0180217

Liao, K.P., Kurreeman, F., Li, G., Duclos, G., Murphy, S., Guzman, R., Cai, T., Gupta, N., Gainer, V., Schur, P., Cui, J., Denny, J.C., Szolovits, P., Churchill, S., Kohane, I., Karlson, E.W., Plenge, R.M., 2013. Associations of autoantibodies, autoimmune risk alleles, and clinical diagnoses from the electronic medical records in rheumatoid arthritis cases and non-rheumatoid arthritis controls. Arthritis & Rheumatism 65, 571–581. https://doi.org/10.1002/art.37801

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, Jill P., Tamayo, P., 2015. The Molecular Signatures Database Hallmark Gene Set Collection. Cell Systems 1, 417–425. https://doi.org/10.1016/j.cels.2015.12.004

Lim, C.P., Phan, T.-T., Lim, I.J., Cao, X., 2006. Stat3 contributes to keloid pathogenesis via promoting collagen production, cell proliferation and migration. Oncogene 25, 5416–5425. https://doi.org/10.1038/sj.onc.1209531

Lim, E.T., Würtz, P., Havulinna, A.S., Palta, P., Tukiainen, T., Rehnström, K., Esko, T., Mägi, R., Inouye, M., Lappalainen, T., Chan, Y., Salem, R.M., Lek, M., Flannick, J., Sim, X., Manning, A., Ladenvall, C., Bumpstead, S., Hämäläinen, E., Aalto, K., Maksimow, M., Salmi, M., Blankenberg, S., Ardissino, D., Shah, S., Horne, B., McPherson, R., Hovingh, G.K., Reilly, M.P., Watkins, H., Goel, A., Farrall, M., Girelli, D., Reiner, A.P., Stitziel, N.O., Kathiresan, S., Gabriel, S., Barrett, J.C., Lehtimäki, T., Laakso, M., Groop, L., Kaprio, J., Perola, M., McCarthy, M.I., Boehnke, M., Altshuler, D.M., Lindgren, C.M., Hirschhorn, J.N., Metspalu, A., Freimer, N.B., Zeller, T., Jalkanen, S., Koskinen, S., Raitakari, O., Durbin, R., MacArthur, D.G., Salomaa, V., Ripatti, S., Daly, M.J., Palotie, A., 2014. Distribution and Medical Impact of Loss-of-Function Variants in the Finnish Founder Population. PLoS Genetics 10, e1004494. https://doi.org/10.1371/journal.pgen.1004494

Lim, I.J., Phan, T.-T., Bay, B.-H., Qi, R., Huynh, H., Tan, W.T.-L., Lee, S.-T., Longaker, M.T., 2002. Fibroblasts cocultured with keloid keratinocytes: normal fibroblasts secrete collagen in a keloidlike manner. American Journal

of Physiology-Cell Physiology 283, C212–C222. https://doi.org/10.1152/ajpcell.00555.2001

Limandjaja, G.C., Broek, Lenie J. van den, Breetveld, M., Waaijman, T., Monstrey, S., Boer, E.M. de, Scheper, R.J., Niessen, F.B., Gibbs, S., 2018a. Characterization of *In Vitro* Reconstructed Human Normotrophic, Hypertrophic, and Keloid Scar Models. Tissue Engineering Part C: Methods 24, 242–253. https://doi.org/10.1089/ten.tec.2017.0464

Limandjaja, G.C., Broek, Leonarda J. van den, Waaijman, T., Breetveld, M., Monstrey, S., Scheper, R.J., Niessen, F.B., Gibbs, S., 2018b. Reconstructed human keloid models show heterogeneity within keloid scars. Archives of Dermatological Research 310, 815–826. https://doi.org/10.1007/s00403-018-1873-1

Limandjaja, G.C., Niessen, F.B., Scheper, R.J., Gibbs, S., 2020b. The keloid disorder: Heterogeneity, histopathology, mechanisms and models. Frontiers in Cell and Developmental Biology 8. https://doi.org/10.3389/fcell.2020.00360

Limandjaja, G.C., Niessen, F.B., Scheper, R.J., Gibbs, S., 2020a. Hypertrophic scars and keloids: Overview of the evidence and practical guide for differentiating between these abnormal scars. Experimental Dermatology 30, 146–161. https://doi.org/10.1111/exd.14121

Limandjaja, G.C., Waaijman, T., Roffel, S., Niessen, F.B., Gibbs, S., 2019. Monocytes co-cultured with reconstructed keloid and normal skin models skew towards M2 macrophage phenotype. Archives of Dermatological Research 311, 615–627. https://doi.org/10.1007/s00403-019-01942-9

Lin, J.-C., Hsiao, W.W.-W., Fan, C.-T., 2020. Transformation of the Taiwan Biobank 3.0: vertical and horizontal integration. Journal of Translational Medicine 18. https://doi.org/10.1186/s12967-020-02451-4

Liu, B., Gloudemans, M.J., Rao, A.S., Ingelsson, E., Montgomery, S.B., 2019. Abundant associations with gene expression complicate GWAS follow-up. Nature Genetics 51, 768–769. https://doi.org/10.1038/s41588-019-0404-0

Liu, J., Ren, J., Su, L., Cheng, S., Zhou, J., Ye, X., Dong, Y., Sun, S., Qi, F., Liu, Z., Pleat, J., Zhai, H., Zhu, N., 2018. Human adipose tissue-derived stem cells inhibit the activity of keloid fibroblasts and fibrosis in a keloid model by paracrine signaling. Burns 44, 370–385. https://doi.org/10.1016/j.burns.2017.08.017

Liu, X., Chen, W., Zeng, Q., Ma, B., Li, Z., Meng, T., Chen, J., Yu, N., Zhou, Z., Long, X., 2022. Single-Cell RNA-Sequencing Reveals Lineage-Specific Regulatory Changes of Fibroblasts and Vascular Endothelial Cells in Keloids. Journal of Investigative Dermatology 142, 124–135.e11. https://doi.org/10.1016/j.

jid.2021.06.010

Liu, Y., Li, Y., Li, N., Teng, W., Wang, M., Zhang, Y., Xiao, Z., 2016. TGF-?1 promotes scar fibroblasts proliferation and transdifferentiation via up-regulating MicroRNA-21. Scientific Reports 6. https://doi.org/10.1038/srep32231

Liu, Y., Wang, X., Yang, D., Xiao, Z., Chen, X., 2014. MicroRNA-21 Affects Proliferation and Apoptosis by Regulating Expression of PTEN in Human Keloid Fibroblasts. Plastic and Reconstructive Surgery 134, 561e–573e. https://doi.org/10.1097/prs.0000000000000577

Lizio, M., Abugessaisa, I., Noguchi, S., Kondo, A., Hasegawa, A., Hon, C.C., de Hoon, M., Severin, J., Oki, S., Hayashizaki, Y., Carninci, P., Kasukawa, T., Kawaji, H., 2018. Update of the FANTOM web resource: expansion to provide additional transcriptome atlases. Nucleic Acids Research 47, D752–D758. https://doi.org/10.1093/nar/gky1099

Lorenz, H.P., Longaker, M.T., Perkocha, L.A., Jennings, R.W., Harrison, M.R., Adzick, N.S., 1992. Scarless wound repair: a human fetal skin model. Development 114, 253–259. https://doi.org/10.1242/dev.114.1.253

Louw, L., 2000. Keloids in rural black South Africans. Part 3: a lipid model for the prevention and treatment of keloid formations. Prostaglandins, Leukotrienes and Essential Fatty Acids (PLEFA) 63, 255–262. https://doi.org/10.1054/plef.2000.0209

Lu, C.-C., Qin, H., Zhang, Z.-H., Zhang, C.-L., Lu, Y.-Y., Wu, C.-H., 2021. The association between keloid and osteoporosis: real-world evidence. BMC Musculoskeletal Disorders 22. https://doi.org/10.1186/s12891-020-03898-8

Lu, W.-S., Zhang, W.-Y., Li, Y., Wang, Z.-X., Zuo, X.-B., Cai, L.-Q., Zhu, F., Wang, J.-F., Sun, L.-D., Zhang, X.-J., Yang, S., 2010. Association of HLA-DRB1 alleles with keloids in Chinese Han individuals. Tissue Antigens 76, 276–281. https://doi.org/10.1111/j.1399-0039.2010.01509.x

Lu, W., Zheng, X., Yao, X., Zhang, L., 2014. Clinical and epidemiological analysis of keloids in Chinese patients. Archives of Dermatological Research 307, 109–114. https://doi.org/10.1007/s00403-014-1507-1

Lu, Y.-Y., Lu, C.-C., Yu, W.-W., Zhang, L., Wang, Q.-R., Zhang, C.-L., Wu, C.-H., 2018. Keloid risk in patients with atopic dermatitis: a nationwide retrospective cohort study in Taiwan. BMJ Open 8, e022865. https://doi.org/10.1136/bmjopen-2018-022865

Lu, Y.-Y., Tu, H.-P., Wu, C.-H., Hong, C.-H., Yang, K.-C., Yang, H.-J., Chang, K.-L., Lee, C.-H., 2021. Risk of cancer development in patients with keloids. Scientific Reports 11. https://doi.org/10.1038/s41598-021-88789-1

Luo, L., Li, J., Liu, H., Jian, X., Zou, Q., Zhao, Q., Le, Q., Chen, H., Gao, X., He, C., 2017. Adiponectin Is Involved in Connective Tissue Growth Factor-Induced Proliferation, Migration and Overproduction of the Extracellular Matrix in Keloid Fibroblasts. International Journal of Molecular Sciences 18, 1044. https://doi.org/10.3390/ijms18051044

Luo, L., Li, J., Wu, Y., Qiao, J., Fang, H., 2021. Adiponectin, but Not TGF- 1, CTGF, IL-6 or TNF- , May Be a Potential Anti-Inflammation and Anti-Fibrosis Factor in Keloid. Journal of Inflammation Research Volume 14, 907–916. https://doi.org/10.2147/jir.s301971

Luo, S., Benathan, M., Raffoul, W., Panizzon, R.G., Egloff, D.V., 2001. Abnormal Balance between Proliferation and Apoptotic Cell Death in Fibroblasts Derived from Keloid Lesions. Plastic and Reconstructive Surgery 107, 87–96. https://doi.org/10.1097/00006534-200101000-00014

Luo, Y., Liu, M., 2016. Adiponectin: a versatile player of innate immunity. Journal of Molecular Cell Biology 8, 120–128. https://doi.org/10.1093/jmcb/mjw012

Ly, L., Winship, I., 2011. X-linked recessive polyfibromatosis manifesting with spontaneous keloid scars and Dupuytren's contracture. Australasian Journal of Dermatology 53, 148–150. https://doi.org/10.1111/j.1440-0960.2011.00740.x

Lykke-Andersen, S., Jensen, T.H., 2015. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nature Reviews Molecular Cell Biology 16, 665–677. https://doi.org/10.1038/nrm4063

Ma, S., Dalgleish, J., Lee, J., Wang, C., Liu, L., Gill, R., Buxbaum, J.D., Chung, W.K., Aschard, H., Silverman, E.K., Cho, M.H., He, Z., Ionita-Laza, I., 2021. Powerful gene-based testing by integrating long-range chromatin interactions and knockoff genotypes. Proceedings of the National Academy of Sciences 118. https://doi.org/10.1073/pnas.2105191118

MacArthur, D.G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K., Jostins, L., Habegger, L., Pickrell, J.K., Montgomery, S.B., Albers, C.A., Zhang, Z.D., Conrad, D.F., Lunter, G., Zheng, H., Ayub, Q., DePristo, M.A., Banks, E., Hu, M., Handsaker, R.E., Rosenfeld, J.A., Fromer, M., Jin, M., Mu, X.J., Khurana, E., Ye, K., Kay, M., Saunders, G.I., Suner, M.-M., Hunt, T., Barnes, I.H.A., Amid, C., Carvalho-Silva, D.R., Bignell, A.H., Snow, C., Yngvadottir, B., Bumpstead, S., Cooper, D.N., Xue, Y., Romero, I.G., Wang, J., Li, Y., Gibbs, R.A., McCarroll, S.A., Dermitzakis, E.T., Pritchard, J.K., Barrett, J.C., Harrow, J., Hurles, M.E., Gerstein, M.B., Tyler-Smith, C., 2012. A Systematic Survey of Loss-of-Function Variants in Human Protein-Coding

Genes. Science 335, 823–828. https://doi.org/10.1126/science.1215040

Mack, M., 2018. Inflammation and fibrosis. Matrix Biology 68-69, 106–121. https://doi.org/10.1016/j.matbio.2017.11.010

Madeira, F., Park, Y. mi, Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., Lopez, R., 2019. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Research 47, W636–W641. https://doi.org/10.1093/nar/gkz268

Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., Lopez, R., 2022. Search and sequence analysis tools services from EMBL-EBI in 2022. Nucleic Acids Research 50, W276–W279. https://doi.org/10.1093/nar/gkac240

Maeda, D., Kubo, T., Kiya, K., Kawai, K., Matsuzaki, S., Kobayashi, D., Fujiwara, T., Katayama, T., Hosokawa, K., 2019. Periostin is induced by IL-4/IL-13 in dermal fibroblasts and promotes RhoA/ROCK pathway-mediated TGF- 1 secretion in abnormal scar formation. Journal of Plastic Surgery and Hand Surgery 53, 288–294. https://doi.org/10.1080/2000656x.2019.1612752

Mah, W., Jiang, G., Olver, D., Cheung, G., Kim, B., Larjava, H., Häkkinen, L., 2014. Human Gingival Fibroblasts Display a Non-Fibrotic Phenotype Distinct from Skin Fibroblasts in Three-Dimensional Cultures. PLoS ONE 9, e90715. https://doi.org/10.1371/journal.pone.0090715

Maier, R.M., Zhu, Z., Lee, S.H., Trzaskowski, M., Ruderfer, D.M., Stahl, E.A., Ripke, S., Wray, N.R., Yang, J., Visscher, P.M., Robinson, M.R., 2018. Improving genetic prediction by leveraging genetic correlations among human diseases and traits. Nature Communications 9.

Marangoni, R.G., Masui, Y., Fang, F., Korman, B., Lord, G., Lee, J., Lakota, K., Wei, J., Scherer, P.E., Otvos, L., Yamauchi, T., Kubota, N., Kadowaki, T., Asano, Y., Sato, S., Tourtellotte, W.G., Varga, J., 2017. Adiponectin is an endogenous anti-fibrotic mediator and therapeutic target. Scientific Reports 7. https://doi.org/10.1038/s41598-017-04162-1

Marconi, G.D., Fonticoli, L., Rajan, T.S., Pierdomenico, S.D., Trubiani, O., Pizzicannella, J., Diomede, F., 2021. Epithelial-Mesenchymal Transition (EMT): The Type-2 EMT in Wound Healing, Tissue Regeneration and Organ Fibrosis. Cells 10, 1587. https://doi.org/10.3390/cells10071587

Marks, R., Barlow, J.W., Funder, J.W., 1982. Steroid-Induced Vasoconstriction: Glucocorticoid Antagonist Studies. The Journal of Clinical Endocrinology & Metabolism 54, 1075–1077. https://doi.org/10.1210/jcem-54-5-1075

Marneros, A.G., 2019. A Role for the E3 Ubiquitin Ligase NEDD4 in Keloid

Pathogenesis. Journal of Investigative Dermatology 139, 279–280. https://doi.org/10.1016/j.jid.2018.09.007

Marneros, A.G., Norris, J.E.C., Olsen, B.R., Reichenberger, E., 2001. Clinical Genetics of Familial Keloids. Archives of Dermatology 137. https://doi.org/10.1001/archderm.137.11.1429

Marneros, A.G., Norris, J.E.C., Watanabe, S., Reichenberger, E., Olsen, B.R., 2004. Genome Scans Provide Evidence for Keloid Susceptibility Loci on Chromosomes 2q23 and 7p11. Journal of Investigative Dermatology 122, 1126–1132. https://doi.org/10.1111/j.0022-202x.2004.22327.x

Marshall, C.D., Hu, M.S., Leavitt, T., Barnes, L.A., Lorenz, H.P., Longaker, M.T., 2018. Cutaneous Scarring: Basic Science, Current Treatments, and Future Directions. Advances in Wound Care 7, 29–45. https://doi.org/10.1089/wound.2016.0696

Marx, V., 2021. Method of the Year: spatially resolved transcriptomics. Nature Methods 18, 9–14. https://doi.org/10.1038/s41592-020-01033-y

Masuda, T., Low, S.-K., Akiyama, M., Hirata, M., Ueda, Y., Matsuda, K., Kimura, T., Murakami, Y., Kubo, M., Kamatani, Y., Okada, Y., 2019. GWAS of five gynecologic diseases and cross-trait analysis in Japanese. European Journal of Human Genetics 28, 95–107. https://doi.org/10.1038/s41431-019-0495-1

Masur, S.K., S, D.H., T, D.T., I, E., S, P., 1996. Myofibroblasts differentiate from fibroblasts when plated at low density. The Proceedings of the National Academy of Sciences 93, 4219–4223. https://doi.org/10.1073/pnas.93.9.4219

Mateu, R., Živicová, V., Krejčí, E.D., Grim, M., Strnad, H., Vlček, Č., Kolář, M., Lacina, L., Gál, P., Borský, J., Smetana, K., Dvořánková, B., 2016. Functional differences between neonatal and adult fibroblasts and keratinocytes: Donor age affects epithelial-mesenchymal crosstalk in vitro. International Journal of Molecular Medicine 38, 1063–1074. https://doi.org/10.3892/ijmm.2016.2706

Matsumoto, N.M., Aoki, M., Okubo, Y., Kuwahara, K., Eura, S., Dohi, T., Akaishi, S., Ogawa, R., 2020. Gene expression profile of isolated dermal vascular endothelial cells in keloids. Frontiers in Cell and Developmental Biology 8. https://doi.org/10.3389/fcell.2020.00658

Mbatchou, J., Barnard, L., Backman, J., Marcketta, A., Kosmicki, J.A., Ziyatdinov, A., Benner, C., O'Dushlaine, C., Barber, M., Boutkov, B., Habegger, L., Ferreira, M., Baras, A., Reid, J., Abecasis, G., Maxwell, E., Marchini, J., 2021. Computationally efficient whole-genome regression for quantitative and binary traits. Nature Genetics 53, 1097–1103. https://doi.org/10.1038/s41588-021-00870-7

McDowell, I.C., Pai, A.A., Guo, C., Vockley, C.M., Brown, C.D., Reddy, T.E., Engelhardt, B.E., 2016. Many long intergenic non-coding RNAs distally regulate mRNA gene expression levels. bioRxiv 044719. https://doi.org/10.1101/044719

McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., Cunningham, F., 2016. The Ensembl Variant Effect Predictor. Genome Biology 17. https://doi.org/10.1186/s13059-016-0974-4

Meehan, T.F., Conte, N., West, D.B., Jacobsen, J.O., Mason, J., Warren, J., Chen, C.-K., Tudose, I., Relac, M., Matthews, P., Karp, N., Santos, L., Fiegel, T., Ring, N., Westerberg, H., Greenaway, S., Sneddon, D., Morgan, H., Codner, G.F., Stewart, M.E., Brown, J., Horner, N., Haendel, M., Washington, N., Mungall, C.J., Reynolds, C.L., Gallegos, J., Gailus-Durner, V., Sorg, T., Pavlovic, G., Bower, L.R., Moore, M., Morse, I., Gao, X., Tocchini-Valentini, G.P., Obata, Y., Cho, S.Y., Seong, J.K., Seavitt, J., Beaudet, A.L., Dickinson, M.E., Herault, Y., Wurst, W., Angelis, M.H. de, Lloyd, K.C.K., Flenniken, A.M., Nutter, L.M.J., Newbigging, S., McKerlie, C., Justice, M.J., Murray, S.A., Svenson, K.L., Braun, R.E., White, J.K., Bradley, A., Flicek, P., Wells, S., Skarnes, W.C., Adams, D.J., Parkinson, H., Mallon, A.-M., Brown, S.D.M., Smedley, D., 2017. Disease model discovery from 3,328 gene knockouts by The International Mouse Phenotyping Consortium. Nature Genetics 49, 1231–1238. https://doi.org/10.1038/ng.3901

Mehrdadi, P., Kolahdouz Mohammadi, R., Alipoor, E., Eshraghian, M., Esteghamati, A., Hosseinzadeh-Attar, M., 2016. The Effect of Coenzyme Q10 Supplementation on Circulating Levels of Novel Adipokine Adipolin/CTRP12 in Overweight and Obese Patients with Type 2 Diabetes. Experimental and Clinical Endocrinology & Diabetes 125, 156–162. https://doi.org/10.1055/s-0042-110570

Messadi, D., 1998. Effect of TGF-beta1 on PDGF receptors expression in human scar fibroblasts. Frontiers in Bioscience 3, a16–22. https://doi.org/10.2741/a246

Messadi, D.V., Le, A., Berg, S., Jewett, A., Wen, Z., Kelly, P., Bertolami, C.N., 1999. Expression of apoptosis-associated genes by human dermal scar fibroblasts. Wound Repair and Regeneration 7, 511–517. https://doi.org/10.1046/j.1524-475x.1999.00511.x

Mikulec, A.A., Hanasono, M.M., Lum, J., Koch, J.M., 2001. Effect of Tamoxifen on Transforming Growth Factor ?$_1$Production by Keloid and Fetal Fibroblasts. Archives of Facial Plastic Surgery 3, 111–114. https://doi.org/10.1001/archfaci.3.2.111

Monaco, J.L., Lawrence, W.T., 2003. Acute wound healing. Clinics in Plastic Surgery 30, 1–12. https://doi.org/10.1016/s0094-1298(02)00070-6

Moschen, A.R., Tilg, H., Raine, T., 2018. IL-12, IL-23 and IL-17 in IBD: immunobiology and therapeutic targeting. Nature Reviews Gastroenterology & Hepatology 16, 185–196. https://doi.org/10.1038/s41575-018-0084-8

Moustafa, M.F.H., Abdel-Fattah, M.A., Abdel-Fattah, D.C., 1975. Presumptive evidence of the effect of pregnancy estrogens on keloid growth. Plastic and Reconstructive Surgery 56, 450–453.

Mukhopadhyay, A., Tan, E.K.J., Khoo, Y.T.A., Chan, S.Y., Lim, I.J., Phan, T.T., 2005. Conditioned medium from keloid keratinocyte/keloid fibroblast coculture induces contraction of fibroblast-populated collagen lattices. British Journal of Dermatology 152, 639–645. https://doi.org/10.1111/j.1365-2133.2005.06545.x

Murray, J.C., Pollack, S.V., Pinnell, S.R., 1981. Keloids: A review. Journal of the American Academy of Dermatology 4, 461–470. https://doi.org/10.1016/s0190-9622(81)70048-3

Myoken, Y., Kan, M., Sato, G.H., McKeehan, W.L., Sato, J.D., 1990. Bifunctional effects of transforming growth factor-? (TGF-?) on endothelial cell growth correlate with phenotypes of TGF-? binding sites. Experimental Cell Research 191, 299–304. https://doi.org/10.1016/0014-4827(90)90018-6

Nagai, A., Hirata, M., Kamatani, Y., Muto, K., Matsuda, K., Kiyohara, Y., Ninomiya, T., Tamakoshi, A., Yamagata, Z., Mushiroda, T., Murakami, Y., Yuji, K., Furukawa, Y., Zembutsu, H., Tanaka, T., Ohnishi, Y., Nakamura, Y., Kubo, M., Shiono, M., Misumi, K., Kaieda, R., Harada, H., Minami, S., Emi, M., Emoto, N., Daida, H., Miyauchi, K., Murakami, A., Asai, S., Moriyama, M., Takahashi, Y., Fujioka, T., Obara, W., Mori, S., Ito, H., Nagayama, S., Miki, Y., Masumoto, A., Yamada, A., Nishizawa, Y., Kodama, K., Kutsumi, H., Sugimoto, Y., Koretsune, Y., Kusuoka, H., Yanai, H., 2017. Overview of the BioBank Japan Project: Study design and profile. Journal of Epidemiology 27, S2–S8. https://doi.org/10.1016/j.je.2016.12.005

Naitoh, M., Kubota, H., Ikeda, M., Tanaka, T., Shirane, H., Suzuki, S., Nagata, K., 2005. Gene expression in human keloids is altered from dermal to chondrocytic and osteogenic lineage. Genes to Cells 10, 1081–1091. https://doi.org/10.1111/j.1365-2443.2005.00902.x

Nakashima, M., Chung, S., Takahashi, A., Kamatani, N., Kawaguchi, T., Tsunoda, T., Hosono, N., Kubo, M., Nakamura, Y., Zembutsu, H., 2010. A genome-wide association study identifies four susceptibility loci for keloid in the Japanese

population. Nature Genetics 42, 768–771. https://doi.org/10.1038/ng.645

Narasimhan, V.M., Hunt, K.A., Mason, D., Baker, C.L., Karczewski, K.J., Barnes, M.R., Barnett, A.H., Bates, C., Bellary, S., Bockett, N.A., Giorda, K., Griffiths, C.J., Hemingway, H., Jia, Z., Kelly, M.A., Khawaja, H.A., Lek, M., McCarthy, S., McEachan, R., O'Donnell-Luria, A., Paigen, K., Parisinos, C.A., Sheridan, E., Southgate, L., Tee, L., Thomas, M., Xue, Y., Schnall-Levin, M., Petkov, P.M., Tyler-Smith, C., Maher, E.R., Trembath, R.C., MacArthur, D.G., Wright, J., Durbin, R., Heel, D.A. van, 2016a. Health and population effects of rare gene knockouts in adult humans with related parents. Science 352, 474–477. https://doi.org/10.1126/science.aac8624

Narasimhan, V.M., Xue, Y., Tyler-Smith, C., 2016b. Human Knockout Carriers: Dead, Diseased, Healthy, or Improved? Trends in Molecular Medicine 22, 341–351. https://doi.org/10.1016/j.molmed.2016.02.006

Nayak, A., Hicks, A.J., Morris, A.A., 2020. Understanding the Complexity of Heart Failure Risk and Treatment in Black Patients. Circulation: Heart Failure 13. https://doi.org/10.1161/circheartfailure.120.007264

Neale, B.M., Sham, P.C., 2004. The Future of Association Studies: Gene-Based Analysis and Replication. The American Journal of Human Genetics 75, 353–362. https://doi.org/10.1086/423901

Nealelab, 2017. Insights from estimates of SNP-heritability for >2,000 traits and disorders in UK biobank (online, accessed may 20 2021).

Negenborn, V.L., Groen, J.-W., Smit, J.M., Niessen, F.B., Mullender, M.G., 2016. The Use of Autologous Fat Grafting for Treatment of Scar Tissue and Scar-Related Conditions: A Systematic Review. Plastic and Reconstructive Surgery 137, 31e–43e. https://doi.org/10.1097/PRS.0000000000001850

Newton, K.M., Peissig, P.L., Kho, A.N., Bielinski, S.J., Berg, R.L., Choudhary, V., Basford, M., Chute, C.G., Kullo, I.J., Li, R., Pacheco, J.A., Rasmussen, L.V., Spangler, L., Denny, J.C., 2013. Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network. Journal of the American Medical Informatics Association 20, e147–e154. https://doi.org/10.1136/amiajnl-2012-000896

Ng, M., Thakkar, D., Southam, L., Werker, P., Ophoff, R., Becker, K., Nothnagel, M., Franke, A., Nürnberg, P., Espirito-Santo, A.I., Izadi, D., Hennies, H.C., Nanchahal, J., Zeggini, E., Furniss, D., 2017. A Genome-wide Association Study of Dupuytren Disease Reveals 17 Additional Variants Implicated in Fibrosis. The American Journal of Human Genetics 101, 417–427. https://doi.org/10.1016/j.ajhg.2017.08.006

Ng, P.C., 2003. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Research 31, 3812–3814. https://doi.org/10.1093/nar/gkg509

Nguyen, J.K., Austin, E., Huang, A., Mamalis, A., Jagdeo, J., 2019. The IL-4/IL-13 axis in skin fibrosis and scarring: mechanistic concepts and therapeutic targets. Archives of Dermatological Research 312, 81–92. https://doi.org/10.1007/s00403-019-01972-3

Nicholson, K.M., Anderson, N.G., 2002. The protein kinase B/Akt signalling pathway in human malignancy. Cellular Signalling 14, 381–395. https://doi.org/10.1016/s0898-6568(01)00271-6

Noishiki, C., Hayasaka, Y., Ogawa, R., 2019. Sex Differences in Keloidogenesis: An Analysis of 1659 Keloid Patients in Japan. Dermatology and Therapy 9, 747–754. https://doi.org/10.1007/s13555-019-00327-0

Noishiki, C., Takagi, G., Kubota, Y., Ogawa, R., 2017. Endothelial dysfunction may promote keloid growth. Wound Repair and Regeneration 25, 976–983. https://doi.org/10.1111/wrr.12601

Nurden, A.T., 2008. Platelets and wound healing. Frontiers in Bioscience Volume, 3525. https://doi.org/10.2741/2947

Nutten, S., 2015. Atopic Dermatitis: Global Epidemiology and Risk Factors. Annals of Nutrition and Metabolism 66, 8–16. https://doi.org/10.1159/000370220

O'Donoghue, M.L., Braunwald, E., White, H.D., Steen, D.P., Lukas, M.A., Tarka, E., Steg, P.G., Hochman, J.S., Bode, C., Maggioni, A.P., Im, K., Shannon, J.B., Davies, R.Y., Murphy, S.A., Crugnale, S.E., Wiviott, S.D., Bonaca, M.P., Watson, D.F., Weaver, W.D., Serruys, P.W., Cannon, C.P., 2014. Effect of Darapladib on Major Coronary Events After an Acute Coronary Syndrome. JAMA 312, 1006. https://doi.org/10.1001/jama.2014.11061

Ogawa, H., Ohashi, K., Ito, M., Shibata, R., Kanemura, N., Yuasa, D., Kambara, T., Matsuo, K., Hayakawa, S., Hiramatsu-Ito, M., Otaka, N., Kawanishi, H., Yamaguchi, S., Enomoto, T., Abe, T., Kaneko, M., Takefuji, M., Murohara, T., Ouchi, N., 2019. Adipolin/CTRP12 protects against pathological vascular remodelling through suppression of smooth muscle cell growth and macrophage inflammatory response. Cardiovascular Research 116, 237–249. https://doi.org/10.1093/cvr/cvz074

Ogawa, R., 2017. Keloid and Hypertrophic Scars Are the Result of Chronic Inflammation in the Reticular Dermis. International Journal of Molecular Sciences 18, 606. https://doi.org/10.3390/ijms18030606

Ogawa, R., Akaishi, S., 2016. Endothelial dysfunction may play a key role in keloid and hypertrophic scar pathogenesis – Keloids and hypertrophic scars may be

vascular disorders. Medical Hypotheses 96, 51–60. https://doi.org/10.1016/j.mehy.2016.09.024

Ogawa, R., Okai, K., Tokumura, F., Mori, K., Ohmori, Y., Huang, C., Hyakusoku, H., Akaishi, S., 2012. The relationship between skin stretching/contraction and pathologic scarring: The important role of mechanical forces in keloid generation. Wound Repair and Regeneration 20, 149–157. https://doi.org/10.1111/j.1524-475x.2012.00766.x

Oluwasanmi, J.O., 1974. Keloids in the African. Clinics in Plastic Surgery 1, 179–195.

Omidifar, A., Toolabi, K., Rahimipour, A., Emamgholipour, S., Shanaki, M., 2019. The gene expression of CTRP12 but not CTRP13 is upregulated in both visceral and subcutaneous adipose tissue of obese subjects. Diabetes & Metabolic Syndrome: Clinical Research & Reviews 13, 2593–2599. https://doi.org/10.1016/j.dsx.2019.07.027

Omo-Dare, P., 1975. Genetic studies on keloid. Journal of the National Medical Association 67, 428–432.

Ong, C., Khoo, Y., Tan, E., Mukhopadhyay, A., Do, D., Han, H., Lim, I., Phan, T., 2006. Epithelial–mesenchymal interactions in keloid pathogenesis modulate vascular endothelial growth factor expression and secretion. The Journal of Pathology 211, 95–108. https://doi.org/10.1002/path.2081

Onoufriadis, A., Hsu, C.-K., Ainali, C., Ung, C.Y., Rashidghamat, E., Yang, H.-S., Huang, H.-Y., Niazi, U., Tziotzios, C., Yang, J.-C., Nuamah, R., Tang, M.-J., Saxena, A., Rinaldis, E. de, McGrath, J.A., 2018. Time Series Integrative Analysis of RNA Sequencing and MicroRNA Expression Data Reveals Key Biologic Wound Healing Pathways in Keloid-Prone Individuals. Journal of Investigative Dermatology 138, 2690–2693. https://doi.org/10.1016/j.jid.2018.05.017

Oriente, A., Fedarko, N.S., Pacocha, S.E., Huang, S.K., Lichtenstein, L.M., Essayan, D.M., 2000. Interleukin-13 modulates collagen homeostasis in human skin and keloid fibroblasts. The Journal of Pharmacology and Experimental Therapeutics 292, 988–994.

Ouchi, N., Parker, J.L., Lugus, J.J., Walsh, K., 2011. Adipokines in inflammation and metabolic disease. Nature Reviews Immunology 11, 85–97. https://doi.org/10.1038/nri2921

Outani, H., Okada, M., Yamashita, A., Nakagawa, K., Yoshikawa, H., Tsumaki, N., 2013. Direct Induction of Chondrogenic Cells from Human Dermal Fibroblast Culture by Defined Factors. PLoS ONE 8, e77365. https://doi.org/10.1371/journal.pone.0077365

Papier, A., Chalmers, R.J.G., Byrnes, J.A., Goldsmith, L.A., 2004. Framework for improved communication: the Dermatology Lexicon Project. Journal of the American Academy of Dermatology 50, 630–634. https://doi.org/10.1016/s0190-9622(03)01571-8

Pavone, D., Clemenza, S., Sorbi, F., Fambrini, M., Petraglia, F., 2018. Epidemiology and Risk Factors of Uterine Fibroids. Best Practice & Research Clinical Obstetrics & Gynaecology 46, 3–11. https://doi.org/10.1016/j.bpobgyn.2017.09.004

Peacock, E.E., Madden, J.W., Trier, W.C., 1970. Biologic Basis for the Treatment of Keloids and Hypertrophic Scars. Southern Medical Journal 63, 755–760. https://doi.org/10.1097/00007611-197007000-00002

Perdigoto, C., 2017. Dawn of the Human Knockout Project. Nature Reviews Genetics 18, 328–329. https://doi.org/10.1038/nrg.2017.35

Perrot, S., Cohen, M., Barke, A., Korwisi, B., Rief, W., Treede, R.-D., 2019. The IASP classification of chronic pain for ICD-11: chronic secondary musculoskeletal pain. Pain 160, 77–82. https://doi.org/10.1097/j.pain.0000000000001389

Phan, L., Jin, Y., Zhang, H., Qiang, W., Shekhtman, E., Shao, D., Revoe, D., Villamarin, R., Ivanchenko, E., Kimura, M., Wang, Z.Y., Hao, L., Sharopova, N., Bihan, M., Sturcke, A., Lee, M., Popova, N., Wu, W., Bastiani, C., Ward, M., Holmes, J.B., Lyoshin, V., Kaur, K., Moyer, E., Feolo, M., Kattman, B.L., 2020. ALFA: Allele Frequency Aggregator. National Center for Biotechnology Information, U.S. National Library of Medicine.

Philippeos, C., Telerman, S.B., Oulès, B., Pisco, A.O., Shaw, T.J., Elgueta, R., Lombardi, G., Driskell, R.R., Soldin, M., Lynch, M.D., Watt, F.M., 2018. Spatial and Single-Cell Transcriptional Profiling Identifies Functionally Distinct Human Dermal Fibroblast Subpopulations. Journal of Investigative Dermatology 138, 811–825. https://doi.org/10.1016/j.jid.2018.01.016

Pletscher-Frankild, S., Pallejà, A., Tsafou, K., Binder, J.X., Jensen, L.J., 2015. DISEASES: Text mining and data integration of disease–gene associations. Methods 74, 83–89. https://doi.org/10.1016/j.ymeth.2014.11.020

Plikus, M.V., Guerrero-Juarez, C.F., Ito, M., Li, Y.R., Dedhia, P.H., Zheng, Y., Shao, M., Gay, D.L., Ramos, R., Hsi, T.-C., Oh, J.W., Wang, X., Ramirez, A., Konopelski, S.E., Elzein, A., Wang, A., Supapannachart, R.J., Lee, H.-L., Lim, C.H., Nace, A., Guo, A., Treffeisen, E., Andl, T., Ramirez, R.N., Murad, R., Offermanns, S., Metzger, D., Chambon, P., Widgerow, A.D., Tuan, T.-L., Mortazavi, A., Gupta, R.K., Hamilton, B.A., Millar, S.E., Seale, P., Pear, W.S., Lazar, M.A., Cotsarelis, G., 2017. Regeneration of fat cells from myofibroblasts

during wound healing. Science (New York, N.Y.) 355, 748–752. https://doi.org/10.1126/science.aai8792

Pratap, U.P., Sharma, H.R., Mohanty, A., Kale, P., Gopinath, S., Hima, L., Priyanka, H.P., ThyagaRajan, S., 2015. Estrogen upregulates inflammatory signals through NF- B, IFN- , and nitric oxide via Akt/mTOR pathway in the lymph node lymphocytes of middle-aged female rats. International Immunopharmacology 29, 591–598. https://doi.org/10.1016/j.intimp.2015.09.024

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., Bakker, P.I.W. de, Daly, M.J., Sham, P.C., 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. The American Journal of Human Genetics 81, 559–575. https://doi.org/10.1086/519795

Qin, H., Lin, Z., Vásquez, E., Luan, X., Guo, F., Xu, L., 2020. Association between obesity and the risk of uterine fibroids: a systematic review and meta-analysis. Journal of Epidemiology and Community Health. https://doi.org/10.1136/jech-2019-213364

Quan, T., Little, E., Quan, H., Qin, Z., Voorhees, J.J., Fisher, G.J., 2013. Elevated Matrix Metalloproteinases and Collagen Fragmentation in Photodamaged Human Skin: Impact of Altered Extracellular Matrix Microenvironment on Dermal Fibroblast Function. Journal of Investigative Dermatology 133, 1362–1366. https://doi.org/10.1038/jid.2012.509

Quang, D., Chen, Y., Xie, X., 2014. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics 31, 761–763. https://doi.org/10.1093/bioinformatics/btu703

Quinlan, A.R., 2014. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. Current Protocols in Bioinformatics 47. https://doi.org/10.1002/0471250953.bi1112s47

Ramakrishnan, K.M., Thomas, K.P., Sundararajan, C.R., 1974. Study of 1,000 patients with keloids in South India. Plastic and Reconstructive Surgery 53, 276–280. https://doi.org/10.1097/00006534-197403000-00004

Rankinen, T., Zuberi, A., Chagnon, Y.C., Weisnagel, S.J., Argyropoulos, G., Walts, B., Pérusse, L., Bouchard, C., 2006. The human obesity gene map: The 2005 update. Obesity 14, 529–644. https://doi.org/10.1038/oby.2006.71

Rausell, A., Luo, Y., Lopez, M., Seeleuthner, Y., Rapaport, F., Favier, A., Stenson, P.D., Cooper, D.N., Patin, E., Casanova, J.-L., Quintana-Murci, L., Abel, L., 2020. Common homozygosity for predicted loss-of-function variants reveals both redundant and advantageous effects of dispensable human genes. Proceedings

of the National Academy of Sciences 117, 13626–13636. https://doi.org/10.1073/pnas.1917993117

Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., Clevers, H., Deplancke, B., Dunham, I., Eberwine, J., Eils, R., Enard, W., Farmer, A., Fugger, L., Göttgens, B., Hacohen, N., Haniffa, M., Hemberg, M., Kim, S., Klenerman, P., Kriegstein, A., Lein, E., Linnarsson, S., Lundberg, E., Lundeberg, J., Majumder, P., Marioni, J.C., Merad, M., Mhlanga, M., Nawijn, M., Netea, M., Nolan, G., Pe'er, D., Phillipakis, A., Ponting, C.P., Quake, S., Reik, W., Rozenblatt-Rosen, O., Sanes, J., Satija, R., Schumacher, T.N., Shalek, A., Shapiro, E., Sharma, P., Shin, J.W., Stegle, O., Stratton, M., Stubbington, M.J.T., Theis, F.J., Uhlen, M., Oudenaarden, A. van, Wagner, A., Watt, F., Weissman, J., Wold, B., Xavier, R., Yosef, N., 2017. The Human Cell Atlas. eLife 6. https://doi.org/10.7554/elife.27041

Reinke, J.M., Sorg, H., 2012. Wound Repair and Regeneration. European Surgical Research 49, 35–43. https://doi.org/10.1159/000339613

Reinke, L., Lam, A.P., Flozak, A.S., Varga, J., Gottardi, C.J., 2016. Adiponectin inhibits Wnt co-receptor, Lrp6, phosphorylation and -catenin signaling. Biochemical and Biophysical Research Communications 470, 606–612. https://doi.org/10.1016/j.bbrc.2016.01.097

Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J., Kircher, M., 2018. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Research 47, D886–D894. https://doi.org/10.1093/nar/gky1016

Ricard-Blum, S., Miele, A.E., 2020. Omic approaches to decipher the molecular mechanisms of fibrosis, and design new anti-fibrotic strategies. Seminars in Cell & Developmental Biology 101, 161–169. https://doi.org/10.1016/j.semcdb.2019.12.009

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H.L., 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine 17, 405–424. https://doi.org/10.1038/gim.2015.30

Rinkevich, Y., Walmsley, G.G., Hu, M.S., Maan, Z.N., Newman, A.M., Drukker, M., Januszyk, M., Krampitz, G.W., Gurtner, G.C., Lorenz, H.P., Weissman, I.L., Longaker, M.T., 2015. Identification and isolation of a dermal lineage with intrinsic fibrogenic potential. Science 348. https://doi.org/10.1126/

science.aaa2151

Ritelli, M., Venturini, M., Cinquina, V., Chiarelli, N., Colombi, M., 2020. Multisystemic manifestations in a cohort of 75 classical Ehlers-Danlos syndrome patients: natural history and nosological perspectives. Orphanet Journal of Rare Diseases 15. https://doi.org/10.1186/s13023-020-01470-0

Rockwell, W.B., Cohen, I.K., Ehrlich, H.P., 1989. Keloids and Hypertrophic Scars. Plastic and Reconstructive Surgery 84, 827–837. https://doi.org/10.1097/00006534-198911000-00021

Romero-Valdovinos, M., Cárdenas-Mejía, A., Gutiérrez-Gómez, C., Flisser, A., Kawa-Karasik, S., Ortiz-Monasterio, F., 2011. Keloid skin scars: the influence of hyperbaric oxygenation on fibroblast growth and on the expression of messenger RNA for insulin like growth factor and for transforming growth factor. In Vitro Cellular & Developmental Biology - Animal 47, 421–424. https://doi.org/10.1007/s11626-011-9418-3

Russell, S.B., Russell, J.D., Trupin, K.M., Gayden, A.E., Opalenik, S.R., Nanney, L.B., Broquist, A.H., Raju, L., Williams, S.M., 2010. Epigenetically Altered Wound Healing in Keloid Fibroblasts. Journal of Investigative Dermatology 130, 2489–2496. https://doi.org/10.1038/jid.2010.162

Russell, S.B., Trupin, K.M., Rodríguez-Eaton, S., Russell, J.D., Trupin, J.S., 1988. Reduced growth-factor requirement of keloid-derived fibroblasts may account for tumor growth. Proceedings of the National Academy of Sciences 85, 587–591. https://doi.org/10.1073/pnas.85.2.587

Rutherford, A., Glass, D.A., 2017. A case-control study analyzing the association of keloids with hypertension and obesity. International Journal of Dermatology 56, e187–e189. https://doi.org/10.1111/ijd.13618

Saito, M., Yamazaki, M., Maeda, T., Matsumura, H., Setoguchi, Y., Tsuboi, R., 2011. Pirfenidone suppresses keloid fibroblast-embedded collagen gel contraction. Archives of Dermatological Research 304, 217–222. https://doi.org/10.1007/s00403-011-1184-2

Sakaue, S., Kanai, M., Tanigawa, Y., Karjalainen, J., Kurki, M., Koshiba, S., Narita, A., Konuma, T., Yamamoto, K., Akiyama, M., Ishigaki, K., Suzuki, A., Suzuki, K., Obara, W., Yamaji, K., Takahashi, K., Asai, S., Takahashi, Y., Suzuki, T., Shinozaki, N., Yamaguchi, H., Minami, S., Murayama, S., Yoshimori, K., Nagayama, S., Obata, D., Higashiyama, M., Masumoto, A., Koretsune, Y., Ito, K., Terao, C., Yamauchi, T., Komuro, I., Kadowaki, T., Tamiya, G., Yamamoto, M., Nakamura, Y., Kubo, M., Murakami, Y., Yamamoto, K., Kamatani, Y., Palotie, A., Rivas, M.A., Daly, M.J., Matsuda, K., Okada, Y., 2021. A cross-

population atlas of genetic associations for 220 human phenotypes. Nature Genetics 53, 1415–1424. https://doi.org/10.1038/s41588-021-00931-x

Saleheen, D., Natarajan, P., Armean, I.M., Zhao, W., Rasheed, A., Khetarpal, S.A., Won, H.-H., Karczewski, K.J., O'Donnell-Luria, A.H., Samocha, K.E., Weisburd, B., Gupta, N., Zaidi, M., Samuel, M., Imran, A., Abbas, S., Majeed, F., Ishaq, M., Akhtar, S., Trindade, K., Mucksavage, M., Qamar, N., Zaman, K.S., Yaqoob, Z., Saghir, T., Rizvi, S.N.H., Memon, A., Hayyat Mallick, N., Ishaq, M., Rasheed, S.Z., Memon, F.-R., Mahmood, K., Ahmed, N., Do, R., Krauss, R.M., MacArthur, D.G., Gabriel, S., Lander, E.S., Daly, M.J., Frossard, P., Danesh, J., Rader, D.J., Kathiresan, S., 2017. Human knockouts and phenotypic analysis in a cohort with a high rate of consanguinity. Nature 544, 235–239. https://doi.org/10.1038/nature22034

Samuels, P., Tan, A.K., 1999. Fetal scarless wound healing. The Journal of Otolaryngology 28, 296–302.

Sargolzaei, J., Chamani, E., Kazemi, T., Fallah, S., Soori, H., 2018. The role of adiponectin and adipolin as anti-inflammatory adipokines in the formation of macrophage foam cells and their association with cardiovascular diseases. Clinical Biochemistry 54, 1–10. https://doi.org/10.1016/j.clinbiochem.2018.02.008

Sattar, N., Nelson, S.M., 2008. Adiponectin, Diabetes, and Coronary Heart Disease in Older Persons: Unraveling the Paradox. The Journal of Clinical Endocrinology & Metabolism 93, 3299–3301. https://doi.org/10.1210/jc.2008-1435

Sazonovs, A., Barrett, J.C., 2018. Rare-Variant Studies to Complement Genome-Wide Association Studies. Annual Review of Genomics and Human Genetics 19, 97–112. https://doi.org/10.1146/annurev-genom-083117-021641

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nature Methods 9, 671–675. https://doi.org/10.1038/nmeth.2089

Scholzen, T., Gerdes, J., 2000. The Ki-67 protein: From the known and the unknown. Journal of Cellular Physiology 182, 311–322. https://doi.org/10.1002/(sici)1097-4652(200003)182:3%3C311::aid-jcp1%3E3.0.co;2-9

Scott, M.A., Nguyen, V.T., Levi, B., James, A.W., 2011. Current Methods of Adipogenic Differentiation of Mesenchymal Stem Cells. Stem Cells and Development 20, 1793–1804. https://doi.org/10.1089/scd.2011.0040

Segura, V., Vilhjálmsson, B.J., Platt, A., Korte, A., Seren, Ü., Long, Q., Nordborg, M., 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nature Genetics 44, 825–830.

https://doi.org/10.1038/ng.2314

Seifert, A.W., Monaghan, J.R., Voss, S.R., Maden, M., 2012. Skin Regeneration in Adult Axolotls: A Blueprint for Scar-Free Healing in Vertebrates. PLoS ONE 7, e32875.

Seifert, O., Mrowietz, U., 2009. Keloid scarring: bench and bedside. Archives of Dermatological Research 301, 259–272. https://doi.org/10.1007/s00403-009-0952-8

Selmanowitz, V.J., 1981. Rubinstein-taybi syndrome. Cutaneous manifestations and colossal keloids. Archives of Dermatology 117, 504–506. https://doi.org/10.1001/archderm.117.8.504

Seoudy, W.M., Mohy El Dien, S.M., Abdel Reheem, T.A., Elfangary, M.M., Erfan, M.A., 2022. Macrophages of the M1 and M2 types play a role in keloids pathogenesis. International Wound Journal. https://doi.org/10.1111/iwj.13834

Sey, N.Y.A., Hu, B., Mah, W., Fauni, H., McAfee, J.C., Rajarajan, P., Brennand, K.J., Akbarian, S., Won, H., 2020. A computational tool (H-MAGMA) for improved prediction of brain-disorder risk genes by incorporating brain chromatin interaction profiles. Nature Neuroscience 23, 583–593. https://doi.org/10.1038/s41593-020-0603-0

Shah, M., Foreman, D.M., Ferguson, M.W., 1995. Neutralisation of TGF-beta 1 and TGF-beta 2 or exogenous addition of TGF-beta 3 to cutaneous rat wounds reduces scarring. Journal of Cell Science 108, 985–1002. https://doi.org/10.1242/jcs.108.3.985

Shaw, T.J., Martin, P., 2009. Wound repair at a glance. Journal of Cell Science 122, 3209–3213. https://doi.org/10.1242/jcs.031187

Shaw, T.J., Rognoni, E., 2020. Dissecting Fibroblast Heterogeneity in Health and Fibrotic Disease. Current Rheumatology Reports 22. https://doi.org/10.1007/s11926-020-00903-w

Shi, H., Gazal, S., Kanai, M., Koch, E.M., Schoech, A.P., Siewert, K.M., Kim, S.S., Luo, Y., Amariuta, T., Huang, H., Okada, Y., Raychaudhuri, S., Sunyaev, S.R., Price, A.L., 2021. Population-specific causal disease effect sizes in functionally important regions impacted by selection. Nature Communications 12. https://doi.org/10.1038/s41467-021-21286-1

Shi, J., Wang, H., Guan, H., Shi, S., Li, Y., Wu, X., Li, N., Yang, C., Bai, X., Cai, W., Yang, F., Wang, X., Su, L., Zheng, Z., Hu, D., 2016. IL10 inhibits starvation-induced autophagy in hypertrophic scar fibroblasts via cross talk between the IL10-IL10R-STAT3 and IL10-AKT-mTOR pathways. Cell Death & Disease 7, e2133–e2133. https://doi.org/10.1038/cddis.2016.44

Shibata, S., Tada, Y., Asano, Y., Hau, C.S., Kato, T., Saeki, H., Yamauchi, T.,

Kubota, N., Kadowaki, T., Sato, S., 2012. Adiponectin Regulates Cutaneous Wound Healing by Promoting Keratinocyte Proliferation and Migration via the ERK Signaling Pathway. The Journal of Immunology 189, 3231–3241. https://doi.org/10.4049/jimmunol.1101739

Shih, B., Bayat, A., 2010. Genetics of keloid scarring. Archives of Dermatological Research 302, 319–339. https://doi.org/10.1007/s00403-009-1014-y

Shilpashree, P., Jaiswal, A.K., Kharge, P.M., 2015. Keloids: An unwanted spontaneity in rubinstein-taybi syndrome. Indian Journal of Dermatology 60, 214. https://doi.org/10.4103/0019-5154.152594

Shim, J., Oh, S.J., Yeo, E., Park, J.H., Bae, J.H., Kim, S.-H., Lee, D., Lee, J.H., 2022. Integrated Analysis of Single-Cell and Spatial Transcriptomics in Keloids: Highlights on Fibrovascular Interactions in Keloid Pathogenesis. Journal of Investigative Dermatology. https://doi.org/10.1016/j.jid.2022.01.017

Shirai, Y., Nakanishi, Y., Suzuki, A., Konaka, H., Nishikawa, R., Sonehara, K., Namba, S., Tanaka, H., Masuda, T., Yaga, M., Satoh, S., Izumi, M., Mizuno, Y., Jo, T., Maeda, Y., Nii, T., Oguro-Igashira, E., Morisaki, T., Kamatani, Y., Nakayamada, S., Nishigori, C., Tanaka, Y., Takeda, Y., Yamamoto, K., Kumanogoh, A., Okada, Y., 2022. Multi-trait and cross-population genome-wide association studies across autoimmune and allergic diseases identify shared and distinct genetic component. Annals of the Rheumatic Diseases. https://doi.org/10.1136/annrheumdis-2022-222460

Shook, B.A., Wasko, R.R., Rivera-Gonzalez, G.C., Salazar-Gatzimas, E., López-Giráldez, F., Dash, B.C., Muñoz-Rojas, A.R., Aultman, K.D., Zwick, R.K., Lei, V., Arbiser, J.L., Miller-Jensen, K., Clark, D.A., Hsia, H.C., Horsley, V., 2018. Myofibroblast proliferation and heterogeneity are supported by macrophages during skin repair. Science 362. https://doi.org/10.1126/science.aar2971

Shumar, S.A., Kerr, E.W., Geldenhuys, W.J., Montgomery, G.E., Fagone, P., Thirawatananond, P., Saavedra, H., Gabelli, S.B., Leonardi, R., 2018. Nudt19 is a renal CoA diphosphohydrolase with biochemical and regulatory properties that are distinct from the hepatic Nudt7 isoform. Journal of Biological Chemistry 293, 4134–4148. https://doi.org/10.1074/jbc.ra117.001358

Sinha, A., Rahman, H., Webb, A., Shah, A.M., Perera, D., 2021. Untangling the pathophysiologic link between coronary microvascular dysfunction and heart failure with preserved ejection fraction. European Heart Journal 42, 4431–4441. https://doi.org/10.1093/eurheartj/ehab653

Sjoberg, D.D., Whiting, K., Curry, M., Lavery, J.A., Larmarange, J., 2021. Reproducible summary tables with the gtsummary package 13, 570–580. https:

//doi.org/10.32614/RJ-2021-053

Slemp, A.E., Kirschner, R.E., 2006. Keloids and scars: a review of keloids and scars, their pathogenesis, risk factors, and management. Current Opinion in Pediatrics 18, 396–402. https://doi.org/10.1097/01.mop.0000236389.41462.ef

Sohl, E., Heymans, M.W., Jongh, R.T. de, Heijer, M. den, Visser, M., Merlijn, T., Lips, P., Schoor, N.M. van, 2014. Prediction of vitamin D deficiency by simple patient characteristics. The American Journal of Clinical Nutrition 99, 1089–1095. https://doi.org/10.3945/ajcn.113.076430

Son, E.D., Lee, J.Y., Lee, S., Kim, M.S., Lee, B.G., Chang, I.S., Chung, J.H., 2005. Topical Application of 17 -Estradiol Increases Extracellular Matrix Protein Synthesis by Stimulating TGF- Signaling in Aged Human Skin In Vivo. Journal of Investigative Dermatology 124, 1149–1161. https://doi.org/10.1111/j.0022-202x.2005.23736.x

St. Laurent, G., Seilheimer, B., Tackett, M., Zhou, J., Shtokalo, D., Vyatkin, Y., Ri, M., Toma, I., Jones, D., McCaffrey, T.A., 2017. Deep sequencing transcriptome analysis of murine wound healing: Effects of a multicomponent, multitarget natural product therapy-Tr14. Frontiers in Molecular Biosciences 4. https://doi.org/10.3389/fmolb.2017.00057

STABILITY Investigators, White, H.D., Held, C., Stewart, R., Tarka, E., Brown, R., Davies, R.Y., Budaj, A., Harrington, R.A., Steg, P.G., Ardissino, D., Armstrong, P.W., Avezum, A., Aylward, P.E., Bryce, A., Chen, H., Chen, M.-F., Corbalan, R., Dalby, A.J., Danchin, N., De Winter, R.J., Denchev, S., Diaz, R., Elisaf, M., Flather, M.D., Goudev, A.R., Granger, C.B., Grinfeld, L., Hochman, J.S., Husted, S., Kim, H.-S., Koenig, W., Linhart, A., Lonn, E., López-Sendón, J., Manolis, A.J., Mohler, E.R.3rd., Nicolau, J.C., Pais, P., Parkhomenko, A., Pedersen, T.R., Pella, D., Ramos-Corrales, M.A., Ruda, M., Sereg, M., Siddique, S., Sinnaeve, P., Smith, P., Sritara, P., Swart, H.P., Sy, R.G., Teramoto, T., Tse, H.-F., Watson, D., Weaver, W.D., Weiss, R., Viigimaa, M., Vinereanu, D., Zhu, J., Cannon, C.P., Wallentin, L., 2014. Darapladib for Preventing Ischemic Events in Stable Coronary Heart Disease. New England Journal of Medicine 370, 1702–1711. https://doi.org/10.1056/nejmoa1315878

Stafford, D.A., Dichmann, D.S., Chang, J.K., Harland, R.M., 2016. Deletion of the sclerotome-enriched lncRNA *PEAT* augments ribosomal protein expression. Proceedings of the National Academy of Sciences 114, 101–106. https://doi.org/10.1073/pnas.1612069113

Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Stein, T.I., Nudel, R., Lieder, I., Mazor, Y., Kaplan, S., Dahary, D., War-

shawsky, D., Guan-Golan, Y., Kohn, A., Rappaport, N., Safran, M., Lancet, D., 2016. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. Current Protocols in Bioinformatics 54. https://doi.org/10.1002/cpbi.5

Stevenson, A.W., Deng, Z., Allahham, A., Prêle, C.M., Wood, F.M., Fear, M.W., 2021. The epigenetics of keloids. Experimental Dermatology 30, 1099–1114. https://doi.org/10.1111/exd.14414

Stewart, E., Cookson, C., Gandolfo, R., Schulze-Rath, R., 2017. Epidemiology of uterine fibroids: a systematic review. BJOG: An International Journal of Obstetrics & Gynaecology 124, 1501–1512. https://doi.org/10.1111/1471-0528.14640

Stirling, D.R., Carpenter, A.E., Cimini, B.A., 2021. CellProfiler Analyst 3.0: accessible data exploration and machine learning for image analysis. Bioinformatics 37, 3992–3994. https://doi.org/10.1093/bioinformatics/btab634

Stone, R.C., Pastar, I., Ojeh, N., Chen, V., Liu, S., Garzon, K.I., Tomic-Canic, M., 2016. Epithelial-mesenchymal transition in tissue repair and fibrosis. Cell and Tissue Research 365, 495–506. https://doi.org/10.1007/s00441-016-2464-0

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., Collins, R., 2015. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLOS Medicine 12, e1001779. https://doi.org/10.1371/journal.pmed.1001779

Sulem, P., Helgason, H., Oddson, A., Stefansson, H., Gudjonsson, S.A., Zink, F., Hjartarson, E., Sigurdsson, G.T., Jonasdottir, A., Jonasdottir, A., Sigurdsson, A., Magnusson, O.T., Kong, A., Helgason, A., Holm, H., Thorsteinsdottir, U., Masson, G., Gudbjartsson, D.F., Stefansson, K., 2015. Identification of a large set of rare complete human knockouts. Nature Genetics 47, 448–452. https://doi.org/10.1038/ng.3243

Sun, K., Tordjman, J., Clément, K., Scherer, Philipp E., 2013. Fibrosis and Adipose Tissue Dysfunction. Cell Metabolism 18, 470–477. https://doi.org/10.1016/j.cmet.2013.06.016

Sun, L.-M., Wang, K.-H., Lee, Y.-C.G., 2014. Keloid incidence in Asian people and its comorbidity with other fibrosis-related diseases: a nationwide population-based study. Archives of Dermatological Research 306, 803–808. https://doi.org/10.1007/s00403-014-1491-5

Sun, X.-J., Wang, Q., Guo, B., Liu, X.-Y., Wang, B., 2017. Identification of

skin-related lncRNAs as potential biomarkers that are involved in Wnt pathways in keloids. Oncotarget 8, 34236–34244. https://doi.org/10.18632/oncotarget.15880

Suttho, D., Mankhetkorn, S., Binda, D., Pazart, L., Humbert, P., Rolin, G., 2016. 3D modeling of keloid scars in vitro by cell and tissue engineering. Archives of Dermatological Research 309, 55–62.

Syed, F., Sherris, D., Paus, R., Varmeh, S., Pandolfi, P.P., Bayat, A., 2012. Keloid Disease Can Be Inhibited by Antagonizing Excessive mTOR Signaling With a Novel Dual TORC1/2 Inhibitor. The American Journal of Pathology 181, 1642–1658. https://doi.org/10.1016/j.ajpath.2012.08.006

Szustakowski, J.D., Balasubramanian, S., Kvikstad, E., Khalid, S., Bronson, P.G., Sasson, A., Wong, E., Liu, D., Wade Davis, J., Haefliger, C., Katrina Loomis, A., Mikkilineni, R., Noh, H.J., Wadhawan, S., Bai, X., Hawes, A., Krasheninina, O., Ulloa, R., Lopez, A.E., Smith, E.N., Waring, J.F., Whelan, C.D., Tsai, E.A., Overton, J.D., Salerno, W.J., Jacob, H., Szalma, S., Runz, H., Hinkle, G., Nioi, P., Petrovski, S., Miller, M.R., Baras, A., Mitnaul, L.J., Reid, J.G., Moiseyenko, O., Rios, C., Saha, S., Abecasis, G., Banerjee, N., Beechert, C., Boutkov, B., Cantor, M., Coppola, G., Economides, A., Eom, G., Forsythe, C., Fuller, E.D., Gu, Z., Habegger, L., Jones, M.B., Lanche, R., Lattari, M., LeBlanc, M., Li, D., Lotta, L.A., Manoochehri, K., Mansfield, A.J., Maxwell, E.K., Mighty, J., Nafde, M., O'Keeffe, S., Orelus, M., Padilla, M.S., Panea, R., Polanco, T., Pradhan, M., Rasool, A., Schleicher, T.D., Sharma, D., Shuldiner, A., Staples, J.C., Van Hout, C.V., Widom, L., Wolf, S.E., John, S., Chen, C.-Y., Sexton, D., Kupelian, V., Marshall, E., Swan, T., Eaton, S., Liu, J.Z., Loomis, S., Jensen, M., Duraisamy, S., Tetrault, J., Merberg, D., Badola, S., Reppell, M., Grundstad, J., Zheng, X., Deaton, A.M., Parker, M.M., Ward, L.D., Flynn-Carroll, A.O., Austin, C., March, R., Pangalos, M.N., Platt, A., Snowden, M., Matakidou, A., Wasilewski, S., Wang, Q., Deevi, S., Carss, K., Smith, K., Sogaard, M., Hu, X., Chen, X., Ye, Z., 2021. Advancing human genetics research and drug discovery through exome sequencing of the UK Biobank. Nature Genetics 53, 942–948. https://doi.org/10.1038/s41588-021-00885-0

Tabib, T., Morse, C., Wang, T., Chen, W., Lafyatis, R., 2018. SFRP2/DPP4 and FMO1/LSP1 Define Major Fibroblast Populations in Human Skin. Journal of Investigative Dermatology 138, 802–810. https://doi.org/10.1016/j.jid.2017.09.045

Tachi, M., Iwamori, M., 2008. Mass spectrometric characterization of cholesterol esters and wax esters in epidermis of fetal, adult and keloidal human skin. Experimental Dermatology 17, 318–323. https://doi.org/10.1111/j.1600-0625.

2007.00647.x

Takikawa, T., Ohashi, K., Ogawa, H., Otaka, N., Kawanishi, H., Fang, L., Ozaki, Y., Eguchi, S., Tatsumi, M., Takefuji, M., Murohara, T., Ouchi, N., 2020. Adipolin/C1q/Tnf-related protein 12 prevents adverse cardiac remodeling after myocardial infarction. PLOS ONE 15, e0243483. https://doi.org/10.1371/journal.pone.0243483

Tal, R., Segars, J.H., 2013. The role of angiogenic factors in fibroid pathogenesis: potential implications for future therapy. Human Reproduction Update 20, 194–216. https://doi.org/10.1093/humupd/dmt042

Tan, B.K., Chen, J., Adya, R., Ramanjaneya, M., Patel, V., Randeva, H.S., 2013. Metformin increases the novel adipokine adipolin/CTRP12: Role of the AMPK pathway. Journal of Endocrinology 219, 101–108. https://doi.org/10.1530/joe-13-0277

Tan, B.K., Chen, J., Hu, J., Amar, O., Mattu, H.S., Ramanjaneya, M., Patel, V., Lehnert, H., Randeva, H.S., 2014a. Circulatory changes of the novel adipokine adipolin/CTRP12 in response to metformin treatment and an oral glucose challenge in humans. Clinical Endocrinology 81, 841–846. https://doi.org/10.1111/cen.12438

Tan, B.K., Lewandowski, K.C., O'Hare, J.P., Randeva, H.S., 2014b. Insulin regulates the novel adipokine adipolin/CTRP12: In vivo and ex vivo effects. Journal of Endocrinology 221, 111–119. https://doi.org/10.1530/joe-13-0537

Tan, S.Y., Lei, X., Little, H.C., Rodriguez, S., Sarver, D.C., Cao, X., Wong, G.W., 2020. CTRP12 ablation differentially affects energy expenditure, body weight, and insulin sensitivity in male and female mice. American Journal of Physiology-Endocrinology and Metabolism 319, E146–E162. https://doi.org/10.1152/ajpendo.00533.2019

Tang, R., Wang, Y.-C., Mei, X., Shi, N., Sun, C., Ran, R., Zhang, G., Li, W., Staveley-O'Carroll, K.F., Li, G., Chen, S.-Y., 2020. LncRNA GAS5 attenuates fibroblast activation through inhibiting Smad3 signaling. American Journal of Physiology-Cell Physiology 319, C105–C115. https://doi.org/10.1152/ajpcell.00059.2020

Thulabandu, V., Chen, D., Atit, R.P., 2017. Dermal fibroblast in cutaneous development and healing. WIREs Developmental Biology 7. https://doi.org/10.1002/wdev.307

Tomasek, J.J., Gabbiani, G., Hinz, B., Chaponnier, C., Brown, R.A., 2002. Myofibroblasts and mechano-regulation of connective tissue remodelling. Nature Reviews Molecular Cell Biology 3, 349–363. https://doi.org/10.1038/nrm809

Tonmoy, M.I.Q., Fariha, A., Hami, I., Kar, K., Reza, H.A., Bahadur, N.M., Hossain, M.S., 2022. Computational epigenetic landscape analysis reveals association of CACNA1G-AS1, F11-AS1, NNT-AS1, and MSC-AS1 lncRNAs in prostate cancer progression through aberrant methylation. Scientific Reports 12. https://doi.org/10.1038/s41598-022-13381-0

Tschoÿ, M., Weyer, C., Tataranni, P.A., Devanarayan, V., Ravussin, E., Heiman, M.L., 2001. Circulating Ghrelin Levels Are Decreased in Human Obesity. Diabetes 50, 707–709. https://doi.org/10.2337/diabetes.50.4.707

Tsekouras, A.A., McGeorge, D.D., 1999. Palmar fasciectomy and keloid formation. British Journal of Plastic Surgery 52, 593–594. https://doi.org/10.1054/bjps.1999.3197

Tsujita-Kyutoku, M., Uehara, N., Matsuoka, Y., Kyutoku, S., Ogawa, Y., Tsubura, A., 2005. Comparison of transforming growth factor-beta/Smad signaling between normal dermal fibroblasts and fibroblasts derived from central and peripheral areas of keloid lesions. In Vivo (Athens, Greece) 19, 959–963.

Turer, A.T., Scherer, P.E., 2012. Adiponectin: mechanistic insights and clinical implications. Diabetologia 55, 2319–2326.

Turner, S., 2018. Qqman: An r package for visualizing GWAS results using q-q and manhattan plots. https://doi.org/10.21105/joss.00731

Ueda, K., Furuya, E., Yasuda, Y., Oba, S., Tajima, S., 1999. Keloids Have Continuous High Metabolic Activity. Plastic and Reconstructive Surgery 104, 694–698. https://doi.org/10.1097/00006534-199909010-00012

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C.A.-K., Odeberg, J., Djureinovic, D., Takanen, J.O., Hober, S., Alm, T., Edqvist, P.-H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J.M., Hamsten, M., Feilitzen, K. von, Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., Heijne, G. von, Nielsen, J., Pontén, F., 2015. Tissue-based map of the human proteome. Science 347. https://doi.org/10.1126/science.1260419

Uitto, J., Tirgan, M.H., 2020. Clinical Challenge and Call for Research on Keloid Disorder: Meeting Report from The 3rd International Keloid Research Foundation Symposium, Beijing 2019. Journal of Investigative Dermatology 140, 515–518. https://doi.org/10.1016/j.jid.2019.10.002

Ung, C.Y., Onoufriadis, A., Parsons, M., McGrath, J.A., Shaw, T.J., 2021. Metabolic perturbations in fibrosis disease. The International Journal of Biochemistry & Cell Biology 139, 106073. https://doi.org/10.1016/j.biocel.2021.106073

Van De Water, L., Varney, S., Tomasek, J.J., 2013. Mechanoregulation of the Myofibroblast in Wound Contraction, Scarring, and Fibrosis: Opportunities for New Therapeutic Intervention. Advances in Wound Care 2, 122–141. https://doi.org/10.1089/wound.2012.0393

Van Hout, C.V., Tachmazidou, I., Backman, J.D., Hoffman, J.D., Liu, D., Pandey, A.K., Gonzaga-Jauregui, C., Khalid, S., Ye, B., Banerjee, N., Li, A.H., O'Dushlaine, C., Marcketta, A., Staples, J., Schurmann, C., Hawes, A., Maxwell, E., Barnard, L., Lopez, A., Penn, J., Habegger, L., Blumenfeld, A.L., Bai, X., O'Keeffe, S., Yadav, A., Praveen, K., Jones, M., Salerno, W.J., Chung, W.K., Surakka, I., Willer, C.J., Hveem, K., Leader, J.B., Carey, D.J., Ledbetter, D.H., Cardon, L., Yancopoulos, G.D., Economides, A., Coppola, G., Shuldiner, A.R., Balasubramanian, S., Cantor, M., Nelson, M.R., Whittaker, J., Reid, J.G., Marchini, J., Overton, J.D., Scott, R.A., Abecasis, G.R., Yerges-Armstrong, L., Baras, A., 2020. Exome sequencing and characterization of 49,960 individuals in the UK Biobank. Nature 586, 749–756. https://doi.org/10.1038/s41586-020-2853-0

VanderWeele, T.J., Tchetgen Tchetgen, E.J., Cornelis, M., Kraft, P., 2014. Methodological Challenges in Mendelian Randomization. Epidemiology 25, 427–435. https://doi.org/10.1097/ede.0000000000000081

Varmeh, S., Egia, A., McGrouther, D., Tahan, S.R., Bayat, A., Pandolfi, P.P., 2011. Cellular Senescence as a Possible Mechanism for Halting Progression of Keloid Lesions. Genes & Cancer 2, 1061–1066. https://doi.org/10.1177/1947601912440877

Velez Edwards, D.R., Tsosie, K.S., Williams, S.M., Edwards, T.L., Russell, S.B., 2014. Admixture mapping identifies a locus at 15q21.2–22.3 associated with keloid formation in African Americans. Human Genetics 133, 1513–1523. https://doi.org/10.1007/s00439-014-1490-9

Vinaik, R., Barayan, D., Auger, C., Abdullahi, A., Jeschke, M.G., 2020. Regulation of glycolysis and the Warburg effect in wound healing. JCI Insight 5. https://doi.org/10.1172/jci.insight.138949

Vincent, A.S., Phan, T.T., Mukhopadhyay, A., Lim, H.Y., Halliwell, B., Wong, K.P., 2008. Human Skin Keloid Fibroblasts Display Bioenergetics of Cancer Cells. Journal of Investigative Dermatology 128, 702–709. https://doi.org/10.1038/sj.jid.5701107

Võsa, U., Claringbould, A., Westra, H.-J., Bonder, M.J., Deelen, P., Zeng, B., Kirsten, H., Saha, A., Kreuzhuber, R., Yazar, S., Brugge, H., Oelen, R., Vries, D.H. de, Wijst, M.G.P. van der, Kasela, S., Pervjakova, N., Alves, I., Favé, M.-J., Agbessi, M., Christiansen, M.W., Jansen, R., Seppälä, I., Tong, L., Teumer, A.,

Schramm, K., Hemani, G., Verlouw, J., Yaghootkar, H., Sönmez Flitman, R., Brown, A., Kukushkina, V., Kalnapenkis, A., Rüeger, S., Porcu, E., Kronberg, J., Kettunen, J., Lee, B., Zhang, F., Qi, T., Hernandez, J.A., Arindrarto, W., Beutner, F., 't Hoen, P.A.C., Meurs, J. van, Dongen, J. van, Iterson, M. van, Swertz, M.A., Jan Bonder, M., Dmitrieva, J., Elansary, M., Fairfax, B.P., Georges, M., Heijmans, B.T., Hewitt, A.W., Kähönen, M., Kim, Y., Knight, J.C., Kovacs, P., Krohn, K., Li, S., Loeffler, M., Marigorta, U.M., Mei, H., Momozawa, Y., Müller-Nurasyid, M., Nauck, M., Nivard, M.G., Penninx, B.W.J.H., Pritchard, J.K., Raitakari, O.T., Rotzschke, O., Slagboom, E.P., Stehouwer, C.D.A., Stumvoll, M., Sullivan, P., 't Hoen, P.A.C., Thiery, J., Tönjes, A., Dongen, J. van, Iterson, M. van, Veldink, J.H., Völker, U., Warmerdam, R., Wijmenga, C., Swertz, M., Andiappan, A., Montgomery, G.W., Ripatti, S., Perola, M., Kutalik, Z., Dermitzakis, E., Bergmann, S., Frayling, T., Meurs, J. van, Prokisch, H., Ahsan, H., Pierce, B.L., Lehtimäki, T., Boomsma, D.I., Psaty, B.M., Gharib, S.A., Awadalla, P., Milani, L., Ouwehand, W.H., Downes, K., Stegle, O., Battle, A., Visscher, P.M., Yang, J., Scholz, M., Powell, J., Gibson, G., Esko, T., Franke, L., 2021. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. Nature Genetics 53, 1300–1310. https://doi.org/10.1038/s41588-021-00913-z

Wainberg, M., Sinnott-Armstrong, N., Mancuso, N., Barbeira, A.N., Knowles, D.A., Golan, D., Ermel, R., Ruusalepp, A., Quertermous, T., Hao, K., Björkegren, J.L.M., Im, H.K., Pasaniuc, B., Rivas, M.A., Kundaje, A., 2019. Opportunities and challenges for transcriptome-wide association studies. Nature Genetics 51, 592–599. https://doi.org/10.1038/s41588-019-0385-z

Wakefield, J., 2009. Bayes factors for genome-wide association studies: comparison with *P*-values. Genetic Epidemiology 33, 79–86. https://doi.org/10.1002/gepi.20359

Wang, K., Li, M., Hakonarson, H., 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research 38, e164–e164. https://doi.org/10.1093/nar/gkq603

Wang, K., Zhang, H., Kugathasan, S., Annese, V., Bradfield, J.P., Russell, R.K., Sleiman, P.M.A., Imielinski, M., Glessner, J., Hou, C., Wilson, D.C., Walters, T., Kim, C., Frackelton, E.C., Lionetti, P., Barabino, A., Van Limbergen, J., Guthery, S., Denson, L., Piccoli, D., Li, M., Dubinsky, M., Silverberg, M., Griffiths, A., Grant, S.F.A., Satsangi, J., Baldassano, R., Hakonarson, H., 2009. Diverse Genome-wide Association Studies Associate the IL12/IL23 Pathway with Crohn Disease. The American Journal of Human Genetics 84, 399–405. https://doi.org/10.1016/j.ajhg.2009.01.026

Wang, Q., Wang, P., Qin, Z., Yang, X., Pan, B., Nie, F., Bi, H., 2021. Altered glucose metabolism and cell function in keloid fibroblasts under hypoxia. Redox Biology 38, 101815. https://doi.org/10.1016/j.redox.2020.101815

Wang, W., Li, J., Wang, K., Zhang, Z., Zhang, W., Zhou, G., Cao, Y., Ye, M., Zou, H., Liu, W., 2016a. Induction of predominant tenogenic phenotype in human dermal fibroblasts via synergistic effect of TGF- and elongated cell shape. American Journal of Physiology-Cell Physiology 310, C357–C372. https://doi.org/10.1152/ajpcell.00300.2015

Wang, W., Qu, M., Xu, L., Wu, X., Gao, Z., Gu, T., Zhang, W., Ding, X., Liu, W., Chen, Y.-L., 2016b. Sorafenib exerts an anti-keloid activity by antagonizing TGF- /Smad and MAPK/ERK signaling pathways. Journal of Molecular Medicine 94, 1181–1194. https://doi.org/10.1007/s00109-016-1430-3

Wang, Y., Zhang, J., 2018. Identification of differential expression lncRNAs in gastric cancer using transcriptome sequencing and bioinformatics analyses. Molecular Medicine Reports. https://doi.org/10.3892/mmr.2018.8889

Wang, Z.-C., Zhao, W.-Y., Cao, Y., Liu, Y.-Q., Sun, Q., Shi, P., Cai, J.-Q., Shen, X.Z., Tan, W.-Q., 2020. The roles of inflammation in keloid and hypertrophic scars. Frontiers in Immunology 11. https://doi.org/10.3389/fimmu.2020.603187

Wang, Z.V., Scherer, P.E., 2016. Adiponectin, the past two decades. Journal of Molecular Cell Biology 8, 93–100. https://doi.org/10.1093/jmcb/mjw011

Warner, J.L., Zollanvari, A., Ding, Q., Zhang, P., Snyder, G.M., Alterovitz, G., 2013. Temporal phenome analysis of a large electronic health record cohort enables identification of hospital-acquired complications. Journal of the American Medical Informatics Association 20, e281–e287. https://doi.org/10.1136/amiajnl-2013-001861

Warwick, A., 2022a. Ukbwranglr: Exploring UKB data.

Warwick, A., 2022b. Codemapper: Functions for mapping between clinical coding systems.

Watanabe, K., Taskesen, E., Bochoven, A. van, Posthuma, D., 2017. Functional mapping and annotation of genetic associations with FUMA. Nature Communications 8. https://doi.org/10.1038/s41467-017-01261-5

Wei, W.-Q., Bastarache, L.A., Carroll, R.J., Marlo, J.E., Osterman, T.J., Gamazon, E.R., Cox, N.J., Roden, D.M., Denny, J.C., 2017. Evaluating phecodes, clinical classification software, and ICD-9-CM codes for phenome-wide association studies in the electronic health record. PLOS ONE 12, e0175508. https://doi.org/10.1371/journal.pone.0175508

Wei, Z., Lei, X., Seldin, M.M., Wong, G.W., 2012a. Endopeptidase Cleavage Generates a Functionally Distinct Isoform of C1q/Tumor Necrosis Factor-related Protein-12 (CTRP12) with an Altered Oligomeric State and Signaling Specificity. Journal of Biological Chemistry 287, 35804–35814. https://doi.org/10.1074/jbc.m112.365965

Wei, Z., Peterson, J.M., Lei, X., Cebotaru, L., Wolfgang, M.J., Baldeviano, G.C., Wong, G.W., 2012b. C1q/TNF-related Protein-12 (CTRP12), a Novel Adipokine That Improves Insulin Sensitivity and Glycemic Control in Mouse Models of Obesity and Diabetes. Journal of Biological Chemistry 287, 10301–10315. https://doi.org/10.1074/jbc.m111.303651

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the {tidyverse} 4, 1686. https://doi.org/10.21105/joss.01686

Wijsman, E.M., 2012. The role of large pedigrees in an era of high-throughput sequencing. Human Genetics 131, 1555–1563. https://doi.org/10.1007/s00439-012-1190-2

Willer, C.J., Li, Y., Abecasis, G.R., 2010. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26, 2190–2191. https://doi.org/10.1093/bioinformatics/btq340

Williams, L.M., McCann, F.E., Cabrita, M.A., Layton, T., Cribbs, A., Knezevic, B., Fang, H., Knight, J., Zhang, M., Fischer, R., Bonham, S., Steenbeek, L.M., Yang, N., Sood, M., Bainbridge, C., Warwick, D., Harry, L., Davidson, D., Xie, W., Sundström, M., Feldmann, M., Nanchahal, J., 2020. Identifying collagen VI as a target of fibrotic diseases regulated by CREBBP/EP300. Proceedings of the National Academy of Sciences 117, 20753–20763. https://doi.org/10.1073/pnas.2004281117

Williams, S.K., Ravenell, J., Seyedali, S., Nayef, H., Ogedegbe, G., 2016. Hypertension Treatment in Blacks: Discussion of the U.S. Clinical Practice Guidelines. Progress in Cardiovascular Diseases 59, 282–288. https://doi.org/10.1016/j.pcad.2016.09.004

Wong, A.J.S., Song, E.J., 2021. Dupilumab as an adjuvant treatment for keloid-associated symptoms. JAAD Case Reports 13, 73–74. https://doi.org/10.1016/j.jdcr.2021.04.034

Woolery-Lloyd, H., Berman, B., 2002. A controlled cohort study examining the onset of hypertension in black patients with keloids. European journal of dermatology:

EJD 12, 581–582.

Wu, J., Del Duca, E., Espino, M., Gontzes, A., Cueto, I., Zhang, N., Estrada, Y.D., Pavel, A.B., Krueger, J.G., Guttman-Yassky, E., 2020. RNA sequencing keloid transcriptome associates keloids with Th2, Th1, Th17/Th22, and JAK3-skewing. Frontiers in Immunology 11. https://doi.org/10.3389/fimmu.2020.597741

Wu, J., Fang, L., Cen, Y., Qing, Y., Chen, J., Li, Z., 2019. MiR-21 Regulates Keloid Formation by Downregulating Smad7 via the TGF- /Smad Signaling Pathway. Journal of Burn Care & Research 40, 809–817. https://doi.org/10.1093/jbcr/irz089

Wu, Y., Wang, B., Li, Y.H., Xu, X.G., Luo, Y.J., Chen, J.Z.S., Wei, H.C., Gao, X.H., Chen, H.D., 2012. Meta-analysis demonstrates association between Arg72Pro polymorphism in the P53 gene and susceptibility to keloids in the chinese population. Genetics and Molecular Research 11, 1701–1711. https://doi.org/10.4238/2012.june.29.2

Wu, Y., Zhang, Q., Ann, D.K., Akhondzadeh, A., Duong, H.S., Messadi, D.V., Le, A.D., 2004. Increased vascular endothelial growth factor may account for elevated level of plasminogen activator inhibitor-1 via activating ERK1/2 in keloid fibroblasts. American Journal of Physiology-Cell Physiology 286, C905–C912. https://doi.org/10.1152/ajpcell.00200.2003

Wynn, T., 2008. Cellular and molecular mechanisms of fibrosis. The Journal of Pathology 214, 199–210. https://doi.org/10.1002/path.2277

Xia, W., Kong, W., Wang, Z., Phan, T.-T., Lim, I.J., Longaker, M.T., Yang, G.P., 2007. Increased CCN2 Transcription in Keloid Fibroblasts Requires Cooperativity Between AP-1 and SMAD Binding Sites. Annals of Surgery 246, 886–895. https://doi.org/10.1097/sla.0b013e318070d54f

Xia, W., Longaker, M.T., Yang, G.P., 2006. P38 MAP kinase mediates transforming growth factor-?2 transcription in human keloid fibroblasts. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology 290, R501–R508. https://doi.org/10.1152/ajpregu.00472.2005

Xia, W., Phan, T.-T., Lim, I.J., Longaker, M.T., Yang, G.P., 2004. Complex epithelial-mesenchymal interactions modulate transforming growth factor-beta expression in keloid-derived cells. Wound Repair and Regeneration 12, 546–556. https://doi.org/10.1111/j.1067-1927.2004.012507.x

Xin, Y., Wang, X., Zhu, M., Qu, M., Bogari, M., Lin, L., Mar Aung, Z., Chen, W., Chen, X., Chai, G., Zhang, Y., 2017. Expansion of CD26 positive fibroblast population promotes keloid progression. Experimental Cell Research. https://doi.org/10.1016/j.yexcr.2017.04.021

Xue, H., McCauley, R.L., Zhang, W., 2000. Elevated Interleukin-6 Expression in Keloid Fibroblasts. Journal of Surgical Research 89, 74–77. https://doi.org/10.1006/jsre.1999.5805

Xue, M., Jackson, C.J., 2015. Extracellular Matrix Reorganization During Wound Healing and Its Impact on Abnormal Scarring. Advances in Wound Care 4, 119–136. https://doi.org/10.1089/wound.2013.0485

Yaghootkar, H., Lamina, C., Scott, R.A., Dastani, Z., Hivert, M.-F., Warren, L.L., Stancáková, A., Buxbaum, S.G., Lyytikäinen, L.-P., Henneman, P., Wu, Y., Cheung, C.Y.Y., Pankow, J.S., Jackson, A.U., Gustafsson, S., Zhao, J.H., Ballantyne, C.M., Xie, W., Bergman, R.N., Boehnke, M., Bouazzaoui, F. el, Collins, F.S., Dunn, S.H., Dupuis, J., Forouhi, N.G., Gillson, C., Hattersley, A.T., Hong, J., Kähönen, M., Kuusisto, J., Kedenko, L., Kronenberg, F., Doria, A., Assimes, T.L., Ferrannini, E., Hansen, T., Hao, K., Häring, H., Knowles, J.W., Lindgren, C.M., Nolan, J.J., Paananen, J., Pedersen, O., Quertermous, T., Smith, U., Lehtimäki, T., Liu, C.-T., Loos, R.J.F., McCarthy, M.I., Morris, A.D., Vasan, R.S., Spector, T.D., Teslovich, T.M., Tuomilehto, J., Dijk, K.W. van, Viikari, J.S., Zhu, N., Langenberg, C., Ingelsson, E., Semple, R.K., Sinaiko, A.R., Palmer, C.N.A., Walker, M., Lam, K.S.L., Paulweber, B., Mohlke, K.L., Duijn, C. van, Raitakari, O.T., Bidulescu, A., Wareham, N.J., Laakso, M., Waterworth, D.M., Lawlor, D.A., Meigs, J.B., Richards, J.B., Frayling, T.M., 2013. Mendelian Randomization Studies Do Not Support a Causal Role for Reduced Circulating Adiponectin Levels in Insulin Resistance and Type 2 Diabetes. Diabetes 62, 3589–3598. https://doi.org/10.2337/db13-0128

Yagi, K.I., Dafalla, A.A., Osman, A.A., 1979. Does an immune reaction to sebum in wounds cause keloid scars? Beneficial effect of desensitisation. British Journal of Plastic Surgery 32, 223–225. https://doi.org/10.1016/s0007-1226(79)90037-7

Yan, X., Gao, J., Chen, Y., Song, M., Liu, X., 2007. [Preliminary linkage analysis and mapping of keloid susceptibility locus in a Chinese pedigree]. Zhonghua Zheng Xing Wai Ke Za Zhi = Zhonghua Zhengxing Waike Zazhi = Chinese Journal of Plastic Surgery 23, 32–35.

Yang, A., Chen, J., Zhao, X.-M., 2020. nMAGMA: a network-enhanced method for inferring risk genes from GWAS summary statistics and its application to schizophrenia. Briefings in Bioinformatics 22. https://doi.org/10.1093/bib/bbaa298

Yang, H., Li, Q., Wu, Y., Dong, J., Lao, Y., Ding, Z., Xiao, C., Fu, J., Bai, S., 2020. Long non-coding RNA RP11-400N13.3 promotes the progression of colorectal cancer by regulating the miR-4722-3p/P2RY8 axis. Oncology Reports. https:

//doi.org/10.3892/or.2020.7755

Yang, J., Lee, S.H., Goddard, M.E., Visscher, P.M., 2011. GCTA: A Tool for Genome-wide Complex Trait Analysis. The American Journal of Human Genetics 88, 76–82. https://doi.org/10.1016/j.ajhg.2010.11.011

Yang, S., Li, H., Yao, H., Zhang, Y., Bao, H., Wu, L., Zhang, C., Li, M., Feng, L., Zhang, J., Zheng, Z., Xu, G., Wang, F., 2021. Long noncoding RNA ERLR mediates epithelial-mesenchymal transition of retinal pigment epithelial cells and promotes experimental proliferative vitreoretinopathy. Cell Death & Differentiation 28, 2351–2366. https://doi.org/10.1038/s41418-021-00756-5

Yang, X., Ni, J., Li, Y., Zou, L., Guo, T., Li, Y., Chu, L., Zhu, Z., 2020. LncRNA-RP11 Modulates TGF- 1-Activated Radiation-Induced Lung Injury Through Downregulating microRNA-29a. Dose-Response 18, 155932582094907. https://doi.org/10.1177/1559325820949071

Yao, R.-W., Wang, Y., Chen, L.-L., 2019. Cellular functions of long noncoding RNAs. Nature Cell Biology 21, 542–551.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., Madden, T.L., 2012. Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics 13. https://doi.org/10.1186/1471-2105-13-134

Young, W.G., Worsham, M.J., Joseph, C.L.M., Divine, G.W., Jones, L.R.D., 2014. Incidence of Keloid and Risk Factors Following Head and Neck Surgery. JAMA Facial Plastic Surgery 16, 379–380. https://doi.org/10.1001/jamafacial.2014.113

Yu, D., Shang, Y., Luo, S., Hao, L., 2013. The TaqI Gene Polymorphisms of VDR and the Circulating 1,25-Dihydroxyvitamin D Levels Confer the Risk for the Keloid Scarring in Chinese Cohorts. Cellular Physiology and Biochemistry 32, 39–45. https://doi.org/10.1159/000350121

Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B., Kresovich, S., Buckler, E.S., 2005. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature Genetics 38, 203–208. https://doi.org/10.1038/ng1702

Yuan, F.-L., Sun, Z.-L., Feng, Y., Liu, S.-Y., Du, Y., Yu, S., Yang, M.-L., Lv, G.-Z., 2019. Epithelial–mesenchymal transition in the formation of hypertrophic scars and keloids. Journal of Cellular Physiology 234, 21662–21669. https://doi.org/10.1002/jcp.28830

Yuan, W., Sun, H., Yu, L., 2021. Long non-coding RNA LINC01116 accelerates the progression of keloid formation by regulating miR-203/SMAD5 axis. Burns 47,

665–675. https://doi.org/10.1016/j.burns.2020.07.027

Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., Sandstrom, R., Ma, Z., Davis, C., Pope, B.D., Shen, Y., Pervouchine, D.D., Djebali, S., Thurman, R.E., Kaul, R., Rynes, E., Kirilusha, A., Marinov, G.K., Williams, B.A., Trout, D., Amrhein, H., Fisher-Aylor, K., Antoshechkin, I., DeSalvo, G., See, L.-H., Fastuca, M., Drenkow, J., Zaleski, C., Dobin, A., Prieto, P., Lagarde, J., Bussotti, G., Tanzer, A., Denas, O., Li, K., Bender, M.A., Zhang, M., Byron, R., Groudine, M.T., McCleary, D., Pham, L., Ye, Z., Kuan, S., Edsall, L., Wu, Y.-C., Rasmussen, M.D., Bansal, M.S., Kellis, M., Keller, C.A., Morrissey, C.S., Mishra, T., Jain, D., Dogan, N., Harris, R.S., Cayting, P., Kawli, T., Boyle, A.P., Euskirchen, G., Kundaje, A., Lin, S., Lin, Y., Jansen, C., Malladi, V.S., Cline, M.S., Erickson, D.T., Kirkup, V.M., Learned, K., Sloan, C.A., Rosenbloom, K.R., Lacerda de Sousa, B., Beal, K., Pignatelli, M., Flicek, P., Lian, J., Kahveci, T., Lee, D., James Kent, W., Ramalho Santos, M., Herrero, J., Notredame, C., Johnson, A., Vong, S., Lee, K., Bates, D., Neri, F., Diegel, M., Canfield, T., Sabo, P.J., Wilken, M.S., Reh, T.A., Giste, E., Shafer, A., Kutyavin, T., Haugen, E., Dunn, D., Reynolds, A.P., Neph, S., Humbert, R., Scott Hansen, R., De Bruijn, M., Selleri, L., Rudensky, A., Josefowicz, S., Samstein, R., Eichler, E.E., Orkin, S.H., Levasseur, D., Papayannopoulou, T., Chang, K.-H., Skoultchi, A., Gosh, S., Disteche, C., Treuting, P., Wang, Y., Weiss, M.J., Blobel, G.A., Cao, X., Zhong, S., Wang, T., Good, P.J., Lowdon, R.F., Adams, L.B., Zhou, X.-Q., Pazin, M.J., Feingold, E.A., Wold, B., Taylor, J., Mortazavi, A., Weissman, S.M., Stamatoyannopoulos, J.A., Snyder, M.P., Guigo, R., Gingeras, T.R., Gilbert, D.M., Hardison, R.C., Beer, M.A., Ren, B., 2014. A comparative encyclopedia of DNA elements in the mouse genome. Nature 515, 355–364. https://doi.org/10.1038/nature13992

Zhan, X., Liu, D.J., 2015. SEQMINER: An r-package to facilitate the functional interpretation of sequence-based associations 39, 1242. https://doi.org/10.1002/gepi.21918

Zhang, G.-Y., Wu, L.-C., Liao, T., Chen, G.-C., Chen, Y.-H., Zhao, Y.-X., Chen, S.-Y., Wang, A.-Y., Lin, K., Lin, D.-M., Yang, J.-Q., Gao, W.-Y., Li, Q.-F., 2015. A novel regulatory function for miR-29a in keloid fibrogenesis. Clinical and Experimental Dermatology 41, 341–345. https://doi.org/10.1111/ced.12734

Zhang, J., Liu, C.Y., Wan, Y., Peng, L., Li, W.F., Qiu, J.X., 2016. Long non-coding RNA H19 promotes the proliferation of fibroblasts in keloid scarring. Oncology Letters 12, 2835–2839. https://doi.org/10.3892/ol.2016.4931

Zhang, M.-Z., Dong, X.-H., Guan, E.-L., Si, L.-B., Zhuge, R.-Q., Zhao, P.-X., Zhang, X., Liu, M.-Y., Adzavon, Y.M., Long, X., Qi, Z., Wang, X., 2017. A comparison

of apoptosis levels in keloid tissue, physiological scars and normal skin. American Journal of Translational Research 9, 5548–5557.

Zhang, T., Goodman, M., Zhu, F., Healy, B., Carruthers, R., Chitnis, T., Weiner, H., Cai, T., De Jager, P., Tremlett, H., Xia, Z., 2020. Phenome-wide examination of comorbidity burden and multiple sclerosis disease severity. Neurology - Neuroimmunology Neuroinflammation 7, e864. https://doi.org/10.1212/nxi.0000000000000864

Zhang, Y., Huang, Y.-X., Jin, X., Chen, J., Peng, L., Wang, D.-L., Li, Y., Yao, X.-Y., Liao, J.-Y., He, J.-H., Hu, K., Lu, D., Guo, Y., Yin, D., 2021. Overexpression of lncRNAs with endogenous lengths and functions using a lncRNA delivery system based on transposon. Journal of Nanobiotechnology 19. https://doi.org/10.1186/s12951-021-01044-7

Zhao, M., Mishra, L., Deng, C.-X., 2018. The role of TGF- /SMAD4 signaling in cancer. International Journal of Biological Sciences 14, 111–123. https://doi.org/10.7150/ijbs.23230

Zhao, Y., Shao, F., 2015. The NAIP-NLRC4 inflammasome in innate immune detection of bacterial flagellin and type III secretion apparatus. Immunological Reviews 265, 85–102. https://doi.org/10.1111/imr.12293

Zhou, M.-Q., Jin, E., Wu, J., Ren, F., Yang, Y.-Z., Duan, D.-D., 2020. CTRP12 Ameliorated Lipopolysaccharide-Induced Cardiomyocyte Injury. Chemical and Pharmaceutical Bulletin 68, 133–139. https://doi.org/10.1248/cpb.c19-00646

Zhou, Y., Liang, Z.-S., Jin, Y., Ding, J., Huang, T., Moore, J.H., Zheng, Z.-J., Huang, J., 2022. Shared genetic architecture and causal relationship between asthma and cardiovascular diseases: A large-scale cross-trait analysis. Frontiers in Genetics 12. https://doi.org/10.3389/fgene.2021.775591

Zhu, F., Wu, B., Li, P., Wang, J., Tang, H., Liu, Y., Zuo, X., Cheng, H., Ding, Y., Wang, W., Zhai, Y., Qian, F., Wang, W., Yuan, X., Wang, J., Ha, W., Hou, J., Zhou, F., Wang, Y., Gao, J., Sheng, Y., Sun, L., Liu, J., Yang, S., Zhang, X., 2013. Association Study Confirmed Susceptibility Loci with Keloid in the Chinese Han Population. PLoS ONE 8, e62377. https://doi.org/10.1371/journal.pone.0062377

Zhu, Z., Lee, P.H., Chaffin, M.D., Chung, W., Loh, P.-R., Lu, Q., Christiani, D.C., Liang, L., 2018. A genome-wide cross-trait analysis from UK Biobank highlights the shared genetic architecture of asthma and allergic diseases. Nature Genetics 50, 857–864. https://doi.org/10.1038/s41588-018-0121-0

Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M.R., Powell, J.E., Montgomery, G.W., Goddard, M.E., Wray, N.R., Visscher, P.M., Yang, J., 2016. Integration

of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nature Genetics 48, 481–487. https://doi.org/10.1038/ng.3538

Zhu, Z., Zhu, X., Liu, C.-L., Shi, H., Shen, S., Yang, Y., Hasegawa, K., Camargo, C.A., Liang, L., 2019. Shared genetics of asthma and mental health disorders: a large-scale genome-wide cross-trait analysis. European Respiratory Journal 54, 1901507. https://doi.org/10.1183/13993003.01507-2019

Živicová, V., Lacina, L., Mateu, R., Smetana, K., Kavková, R., Krejčí, E.D., Grim, M., Kvasilová, A., Borský, J., Strnad, H., Hradilová, M., Šáchová, J., Kolář, M., Dvořánková, B., 2017. Analysis of dermal fibroblasts isolated from neonatal and child cleft lip and adult skin: Developmental implications on reconstructive surgery. International Journal of Molecular Medicine 40, 1323–1334. https://doi.org/10.3892/ijmm.2017.3128

Żółkiewicz, J., Stochmal, A., Rudnicka, L., 2019. The role of adipokines in systemic sclerosis: a missing link? Archives of Dermatological Research 311, 251–263. https://doi.org/10.1007/s00403-019-01893-1

Zuffardi, O., Fraccaro, M., 1982. Gene mapping and serendipity. The locus for torticollis, keloids, cryptorchidism and renal dysplasia (31430, McKusick) is at Xq28, distal to the G6PD locus. Human Genetics 62, 280–281. https://doi.org/10.1007/bf00333537