The genetic basis of dermatophytosis skin infection susceptibility

Hele Haapaniemi [1], Reyhane Eghtedarian [1], Anniina Tervi [1], Estonian Biobank Research Team[2], FinnGen, Erik Abner [2], Hanna M. Ollila [1, 3-5]

1) Institute for Molecular Medicine Finland, FIMM, HiLIFE, University of Helsinki, Helsinki, Finland

2) Functional and Population Genomics, Institute of Genomics, University of Tartu, Tartu, Estonia

3) Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA

4) Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

5) Center for Genomic Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

Abstract

Dermatophytosis is an infection caused by fungi that utilize keratinized tissues, such as skin, nails, and hair, as their energy source. This infection commonly presents as red, itchy and ring-like patches on the skin, nail thickening, or hair loss. With ever-increasing case numbers, it has become a significant public health concern estimated to affect 20 % of the world's population. Despite the high prevalence, the genetic risk factors for dermatophytosis are poorly understood. Our goal was to elucidate the biological mechanisms underlying individual susceptibility to dermatophytosis and to explore its genetic associations with other diseases and traits. We performed a large-scale genome-wide association meta-analysis of dermatophytosis infections with over 250,000 cases and 1,370,000 controls using data from FinnGen, Estonian Biobank, UK Biobank and Million Veterans Program. We identified 30 genome-wide significant loci including seven missense variants and two variants in high linkage disequilibrium with missense variants. The strongest associations were with variants within or closest to *ZNF646* ($p = 6.60 \times 10^{-79}$, beta = 0.07), *HLA-DQB1* ($p = 1.42 \times 10^{-36}$, beta = 0.05), *FLG* ($p = 1.96 \times 10^{-27}$, beta = -0.22), *FTO* ($p = 5.75 \times 10^{-26}$, beta = -0.04), *SLURP2* ($p = 3.33 \times 10^{-24}$, beta = 0.04) and *KRT77* ($p = 1.28 \times 10^{-15}$, beta = 0.03) genes. Overall, our findings implicate keratin lifecycle and skin integrity, immune defense, and obesity as risk factors for dermatophytosis. Our findings highlight the clinical comorbidities with other skin diseases and with high BMI and identify novel genetic variants some of which are novel candidates for managing dermatophytosis infection.

Introduction

Dermatophytosis, commonly known as ringworm, is a prevalent fungal infection affecting the skin, hair, and nails. It is caused by dermatophytes, a group of keratinophilic fungi with the unique ability to utilize keratin, a structural protein in the outer layer of human skin, as a nutrient source, leading to various clinical symptoms [1,2]. While typically limited to the outermost layers of the skin, dermatophytosis can become more severe in certain patient groups. For example in immunocompromised and diabetic patients with compromised immune system or skin barrier, the infection may invade deeper layers of the skin and can lead to severe and invasive disease [3].

The global incidence of dermatophytosis has made it a significant public health concern, particularly in regions with warm and humid climates that favor fungal growth. It has been estimated that around 20-25 % of people are infected with dermatophytes at some point in their lives and the incidence rate is constantly rising [4]. The prevalence varies between continents and countries but more recent studies from European countries have estimated prevalence rates around 12-17 % [5,6]. Besides geographic location, the individual vulnerability to dermatophytosis depends on several factors including age, sex, season, socioeconomic status, personal hygiene and cultural conditions [7,8]. In addition, existing skin diseases or skin lesions together with immunocompromising factors can affect the susceptibility to dermatophytosis.

Dermatophytosis presents a wide range of symptoms, depending on the site of infection. On the skin (e.g. tinea corporis, tinea pedis), it typically manifests as red, scaly, and itchy patches that often form a ring-like pattern, hence the name "ringworm". Infections of the scalp (tinea capitis) can lead to hair loss and inflammation, while nail infections (onychomycosis) cause thickening, discoloration, and brittleness of the nails [9]. The infection is highly contagious and can spread through direct contact with infected individuals or animals, as well as indirectly through contaminated objects like clothing, towels, and grooming tools [10].

While environmental factors associate with dermatophytosis infections, genetic studies provide an avenue to understand novel biological mechanisms that contribute to risk and development of dermatophytosis. Here we aimed to understand host factors that affect susceptibility to dermatophytosis infections by performing the largest genome-wide association analysis with over 250,000 dermatophytosis cases from FinnGen, the UK Biobank, the Estonian Biobank and the Million Veteran Program. Our findings highlight the role of barrier organs and variety of immune functions in the development of dermatophytosis.

Results

GWAS shows an association between dermatophytosis and 30 genetic loci

To explore the host genetic components contributing to dermatophytosis, we performed GWAS and meta-analysis in FinnGen (N = 27,662 cases and 471,729 controls), UK Biobank (N = 27,755 cases and 380,368 controls), Estonian biobank (N = 50,241 cases and 106,586 controls) and Million Veterans Program (N = 151,164 cases and 413,818 controls).

With data from 256,822 dermatophytosis cases and 1,372,501 controls, we identified 30 genome-wide significant loci ($p < 5x10^{-8}$) associated with dermatophytosis infection (Figure 1, Table 1, Table S1). The most significant loci were *ZNF646* ($p = 6.60x10^{-79}$), *HLA* ($p = 1.42x10^{-36}$), *FLG* ($p = 1.96x10^{-27}$), *SLURP2* ($p = 3.33x10^{-24}$) and *KRT77* ($p = 1.28x10^{-15}$).

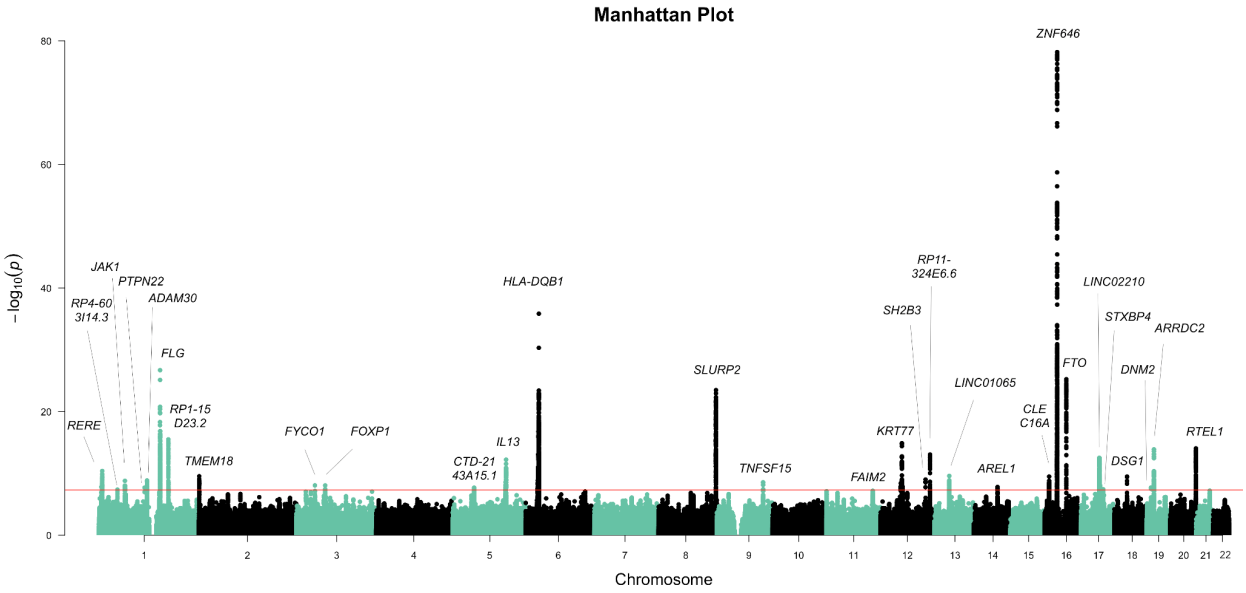

Figure 1. Manhattan plot for Dermatophytosis infection meta-analysis using data from FinnGen, the Estonian Biobank, the UK Biobank and the Million Veterans Program biobank comprising 256,822 cases and 1,372,501 controls. The X-axis represents the chromosomal position for each variant. The Y-axis shows the $-\log_{10}$(P)-value. The horizontal line indicates the genome-wide significance threshold of $p = 5x10^{-8}$.

Table 1. Genome-wide significant lead variants from dermatophytosis infection meta-analysis using data from FinnGen, Estonian biobank, UK biobank and Million Veterans Program biobank comprising 256,822 cases and 1,372,501 controls. Lead variants (SNP), their effect sizes (BETA), p-values (P), allele frequencies (AF), nearest genes, variant types (most severe consequence) and direction of the effect in each cohort (in the following order: UK biobank, FinnGen, Estonian biobank, Million Veterans program; question marks indicating missingness of the SNP in a particular data) are presented.

SNP	CHR	AF	BETA	SE	P	Direction	Nearest gene	Most severe consequence
rs6577497	1	0.640	0.024	0.004	4.05×10^{-11}	++++	*RERE*	intron variant
rs56733721	1	0.605	-0.020	0.004	4.24×10^{-08}	----	*RP4-603I14.3*	downstream gene variant
rs506025	1	0.839	-0.030	0.005	1.58×10^{-09}	----	*JAK1*	intergenic variant
rs2476601	1	0.111	0.032	0.006	2.03×10^{-08}	++++	*PTPN22*	missense variant
rs35273427	1	0.910	-0.038	0.006	1.40×10^{-09}	----	*ADAM30*	missense variant
rs558269137	1	0.984	-0.222	0.021	1.96×10^{-27}	?---	*FLG FLG-AS1*	frameshift variant
rs10912488	1	0.730	0.032	0.004	3.16×10^{-16}	++++	*RP1-15D23.2*	intron variant
rs60383093	2	0.829	0.032	0.005	2.90×10^{-10}	+?++	*TMEM18*	intergenic variant
rs11130078	3	0.409	-0.021	0.004	8.73×10^{-09}	----	*FYCO1*	intron variant
rs62246017	3	0.281	-0.023	0.004	8.86×10^{-09}	----	*FOXP1*	intron variant
rs37807	5	0.684	-0.021	0.004	2.05×10^{-08}	----	*CTD-2143A15.1*	intergenic variant
rs848	5	0.295	-0.028	0.004	5.95×10^{-13}	----	*IL13 TH2LCRR*	3 prime UTR variant
rs1794269	6	0.401	0.046	0.004	1.42×10^{-36}	++++	*HLA-DQB1*	intergenic variant
rs10094888	8	0.362	0.036	0.004	3.33×10^{-24}	++++	*SLURP2*	upstream gene variant
rs56069985	9	0.944	0.049	0.008	2.80×10^{-09}	++++	*TNFSF15*	upstream gene variant
rs640081	12	0.659	0.022	0.107	3.89×10^{-09}	++++	*FAIM2*	intron variant
rs1829637	12	0.303	0.031	0.004	1.28×10^{-15}	++++	*KRT77*	intron variant
rs3184504	12	0.419	0.022	0.004	9.74×10^{-10}	++++	*SH2B3*	missense variant
rs612057	12	0.455	0.026	0.004	8.79×10^{-14}	++++	*RP11-324E6.6*	intron variant

rs1885767	13	0.398	0.023	0.004	$2.56x10^{-10}$	++++	*LINC01065*	intergenic variant
rs12434646	14	0.422	-0.031	0.005	$1.58x10^{-08}$	???-	*AREL1*	intron variant
rs34306440	16	0.804	0.028	0.005	$3.21x10^{-10}$	++++	*CLEC16A*	intron variant
rs7196726	16	0.413	0.066	0.004	$6.60x10^{-79}$	++++	*ZNF646*	missense variant
rs1421085	16	0.616	-0.038	0.004	$5.75x10^{-26}$	----	*FTO*	intron variant
rs62066119	17	0.189	0.039	0.005	$3.05x10^{-13}$	++?+	*LINC02210*	upstream gene variant
rs244304	17	0.708	-0.022	0.004	$3.65x10^{-08}$	----	*STXBP4*	3 prime UTR variant
rs61730306	18	0.913	-0.038	0.006	$3.27x10^{-10}$	----	*DSG1*	missense variant
rs2043332	19	0.292	-0.022	0.004	$2.00x10^{-08}$	----	*DNM2*	intron variant
rs438735	19	0.094	-0.046	0.006	$1.30x10^{-14}$	----	*ARRDC2*	downstream gene variant
rs2236506	20	0.785	-0.034	0.004	$8.96x10^{-15}$	----	*RTEL1 RTEL1-TNFRSF6B*	missense variant

The majority of the genetic associations of complex diseases are regulatory variants typically located at the non-coding or intronic regions of the genome and usually, these associations typically affect gene expression levels rather than have direct impact on protein structure [11]. In our meta-analysis nine out of thirty variants were either missense variants or in high LD (linkage disequilibrium) with a missense variant. (Table 2)

Table 2. Missense lead variants and missense variants in LD with our lead variants reported with the effect gene, amino acid change, p-value, beta and PolyPhen score describing the damaging effect of amino acid change. In the GWAS lead variant column, the variants marked with an asterisk (*) are not missense variants but have missense variants in high LD (d' > 0.9) with them.

CHR	**GWAS lead variant**	**missense variant in high LD**	**gene**	**amino acid change**	**LD**	**P**	**BETA**	**PolyPhen Score**
1	rs2476601	NA	*PTPN22*	Trp620Arg	NA	$2.03x10^{-08}$	0.032	0
1	rs35273427	NA	*ADAM30*	Thr737Ala	NA	$1.40x10^{-09}$	-0.038	0
1	rs558269137	NA	*FLG FLG-AS1*	Ser761CysfsTer36	NA	$1.96x10^{-27}$	-0.222	NA
5	rs848*	rs20541	*IL13, TH2LCRR*	Gln144Arg	d'=1 R^2=0.9660	$1.28x10^{-11}$	-0.027	0
12	rs1829637*	rs14024	*KRT1*	Lys633Arg	d'=0.9957 R^2=0.987	$3.95x10^{-15}$	-0.030	0.078

		rs10783528	*KRT77*	Asp336Asn	d'=1 R^2=0.8876	$9.75x10^{-12}$	0.025	0.039
12	rs3184504	NA	*SH2B3*	Trp262Arg	NA	$9.74x10^{-10}$	0.022	0.039
16	rs7196726	NA	*ZNF646*	Gly1477Asp	NA	$6.60x10^{-79}$	0.066	0.003
		rs35713203	*ZNF646*	Gly921Ala	d'=0.9833 R^2=0.9355	$3.19x10^{-76}$	0.066	0.015
		rs7199949	*PRSS53*	Pro406Ala	d'=1 R^2=0.9918	$1.90x10^{-78}$	0.065	0
		rs9938550	*HSD3B7*	Thr250Ala	d'=0.9119 R^2=0.8013	$3.43x10^{-57}$	0.056	0
18	rs61730306	NA	*DSG1*	Lys537Arg	NA	$3.27x10^{-10}$	-0.038	0
		rs16961689	*DSG1*	Tyr528Ser	d'=1 $R^2$2=1	$3.49x10^{-10}$	-0.038	0
		rs34302455	*DSG1*	Asp538Asn	d'=1 R^2=1	$3.84x10^{-10}$	0.038	0
20	rs2236506	NA	*RTEL1 RTEL1-TNFRSF6B*	Ala758Ala	NA	$8.96x10^{-15}$	-0.034	0.007

The missense variants with the most significant association with dermatophytosis infection were identified within the *ZNF646* gene (rs7196726, beta = 0.066, p = $6.60x10^{-79}$) and its neighboring genes *PRSS53* (rs35713203, beta = 0.065, p = $1.90x10^{-78}$) and *HSD3B7* (rs9938550, beta = 0.056, p = $3.43x10^{-57}$). In addition, we identified highly significant missense variants (< $5x10^{-15}$) within the *FLG* gene (rs558269137, beta = -0.222, p = $1.96x10^{-27}$) and the *KRT1* gene (rs14024, beta = -0.030, p = $3.95x10^{-15}$). All of the reported missense variants have minor allele frequency above 1 % and all of them are predicted to be benign based on their Polyphen score (<0.15), estimating impact of an amino acid substitution on the structure and function of a human protein [12].

While missense variation can indicate a likely causal gene at the locus, the majority of the associated variants were located in non-coding or intronic regions and likely contribute to disease risk by affecting gene expression. To elucidate possible affected genes near the regions of the strongest non-missense associations (rs10094888 closest to *SLURP2* and rs1794269 closest to *FTO*), we performed a colocalization analysis with expression data from GTEx [13] (https://gtexportal.org/home/) (Table S2-S3). We identified a strong shared signal between *SLURP2* (rs10094888) and a structurally and functionally highly similar Ly6SF-group gene *LYNX* expression in skin tissue and in dermatophytosis with posterior probabilities of 0.999 (*LYNX1*) and 0.987 (*SLURP2*). The results indicate the same causal variant for differential expression of *SLURP2* and *LYNX1* in dermatophytosis and skin tissue (Figure 2A-B). These proteins are secreted primarily in the skin by keratinocytes and have been previously implicated in skin diseases [14].

We also discovered associations in canonical loci that have been earlier reported in infections and obesity including the *FTO* locus. The lead variant at the *FTO* locus was the same variant that has been implicated as a causal variant in higher body mass index (rs1421085), and formal colocalization analysis showed a signal with *IRX3*, with a posterior probability of 0.86 (Figure 2C) aligning with the signal reported earlier for obesity [15]. These findings may implicate the role of BMI (body mass index) in dermatophytosis as suggested in previous studies [16].

In addition to showing the likely association of variants from *SLURP2* and *FTO* to dermatophytosis and aforementioned missense variants, our GWAS shows an association between the HLA region and dermatophytosis infections, aligning with previous studies [17,18]. The lead variant for the association (rs1794269) was located closest to the *HLA-DQB1* gene (beta = 0.046 and p = 1.42×10^{-36}). The association between *HLA* region and dermatophytosis highlights an overall immune signal in dermatophytosis.

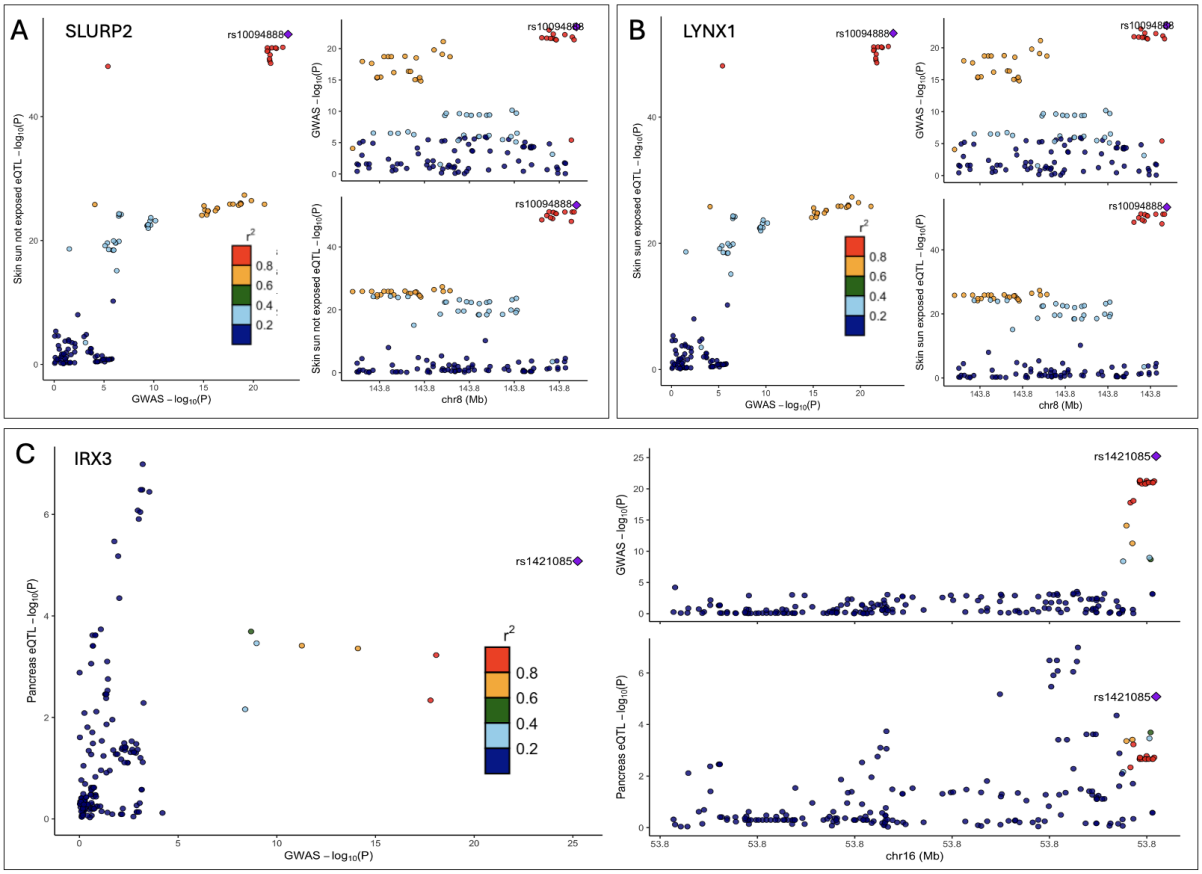

Figure 2. Colocalization plots for Dermatophytosis infection and relevant tissue. A. Colocalization plot for *SLURP2* and not sun-exposed skin tissue. B. Colocalization plot for *LYNX1* and sun-exposed skin tissue. C. Colocalization plot for *IRX3* and pancreatic tissue.

GWAS associations highlight the role of compromised keratin processing behind dermatophytosis

Dermatophytes are fungi with the unique ability to utilize keratin as a source of energy, allowing them to invade and colonize the outer layers of skin, hair, and nails, causing superficial infections. [19] We identified several lead variants that were located within or near genes involved in the keratin lifecycle, including the expression, differentiation, migration, and apoptosis of keratin-producing cells (keratinocytes). The strongest associations were identified in Profilagrin (*FLG)* that forms the outermost layer of the skin, keratin 1 (*KRT1)*, and *SLURP2*, all of which play critical roles in keratin function. These findings highlight keratin's crucial role in the development and progression of dermatophytosis. Many keratin-related proteins are essential for maintaining skin integrity and barrier function, which are key to protecting against fungal infections. [20]

Moreover, we observed the strongest association at *ZNF646*. *ZNF646* belongs to a family of zinc finger proteins that act in transcriptional regulation and protein degradation [21]. This protein family is implicated in tissue development, particularly in the skin where several family members can modulate keratinocyte gene expression and differentiation [21]. Zinc finger proteins have been earlier implicated in the regulation of *FLG* [21,22] This observation raises a possibility that *ZNF646* modulates *FLG* expression. To test this, we examined the effect of the lead missense variant, rs7196726, at *ZNF646* on *FLG* and *KRT1* expression in the skin using GTEx eQLT calculator (https://gtexportal.org/home/testyourown). We observed that *ZNF646* was a trans-eQTL for *FLG* expression in the sun-exposed skin (NES= -0.063 and p = 0.0039) as well as in the not sun-exposed skin (NES= -0.058 and p = 0.031) suggesting an effector role of *ZNF646* on *FLG*.

Another significant group of genes associated with dermatophytosis involves those related to immune defense, such as *HLA* genes, *IL13*, and *SH2B3* (*LNK*). These genes play crucial roles in antigen presentation, immune signaling, and cytokine production, all of which are essential for the body's defense against pathogens [23-25].

A third group of genes identified is related to obesity. Among these, the most notable was the association with *FTO* and *IRX3*, with *IRX3* mediating the functional effects of *FTO* [15]. The structure of the skin, the roles of each layer, and the potential causal genes for dermatophytosis—grouped by their function—are presented in Figure 3.

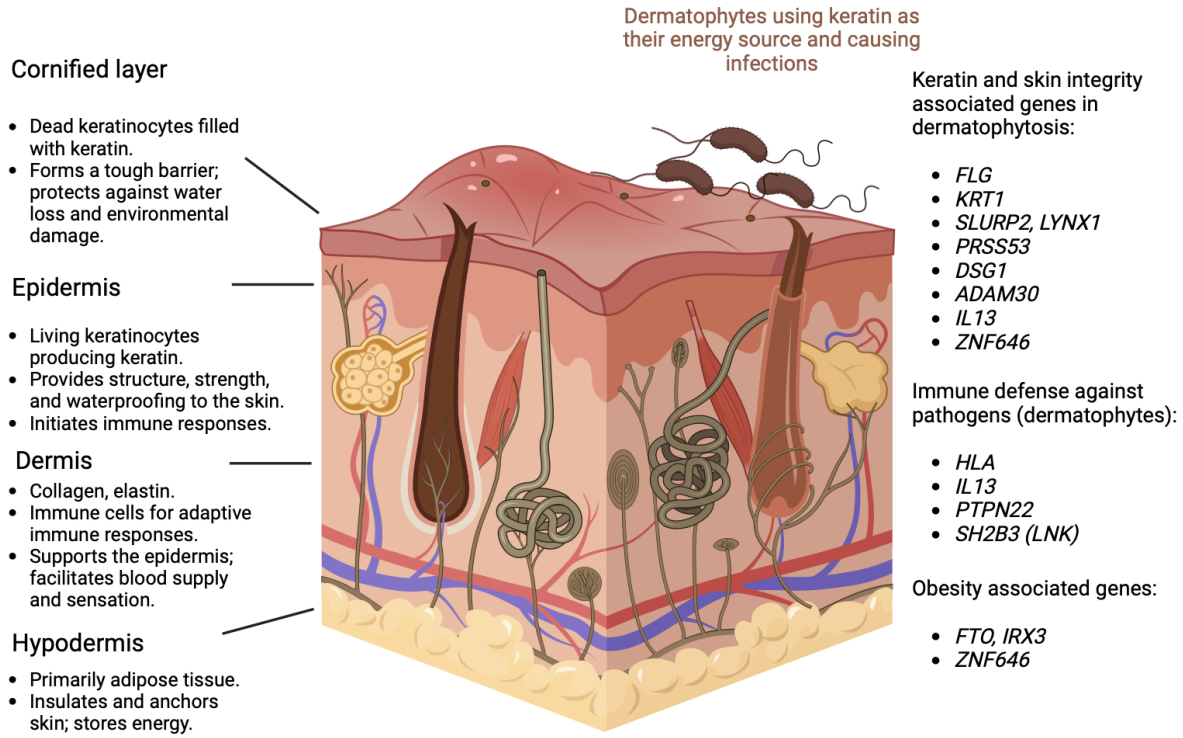

Figure 3. Genes associated with dermatophytosis were associated with keratin processing and skin integrity, systemic immune defense against pathogens or obesity. Picture created with Biorender (https://www.biorender.com/).

Variant annotations support associations with skin well-being, general immune defense and BMI

To get a broader understanding of lead variants in dermatophytosis GWAS, we assessed their associations with other diseases and traits using FinnGen's annotation tool (https://anno.finngen.fi/, https://github.com/juhis/genetics-results-browser) covering all endpoints in FinnGen data freeze 12 (R12) (https://risteys.finregistry.fi/), FinnGen R12 and UKBB combined meta-analysis and Open targets (https://platform.opentargets.org/). The strongest genome-wide significant ($p < 5x10^{-8}$) signal for each variant is shown in Table 3. Besides the strongest associations presented in the table, most of the variants are associated with multiple or up to few hundred other traits at a genome-wide significant level.

Several lead variants are most strongly associated with circulating immune cells and autoimmune diseases, highlighting the connection of dermatophytosis to suboptimal internal immune defense. Secondly, we see associations with vitamin D levels, a vitamin that is essential for skin well-being and is known to regulate, for example, epidermal keratinocytes [26]. A few of the most significant genetic associations are also directly linked with skin diseases such as atopic dermatitis or skin cancer. Lastly, we see a group of associations with traits related to high body weight.

Table 3. Most significant associations for our dermatophytosis meta-analysis GWAS lead variants with endpoints from FinnGen release 12 (R12), FinnGen R12 and UKBB meta-analysis and Open targets.

variant	SNP	nearest gene	top association	P	BETA
1-113834946-A-G	rs2476601	*PTPN22*	Autoimmune diseases	10.00×10^{-403}	-0.25
1-119894128-T-C	rs35273427	*ADAM30*	Type 2 diabetes wide definition	6.31×10^{-14}	0.09
1-152312600-CACTG-C	rs558269137	*FLG*	Vitamin D levels	3.70×10^{-97}	0.15
1-172877652-A-G	rs10912488	*RP1-15D23.2*	Crohn's disease	1.87×10^{-21}	0.17
1-65014876-A-G	rs506025	*JAK1*	Eosinophil counts	4.40×10^{-30}	0.04
1-8545608-A-T	rs6577497	*RERE*	Heel bone mineral density	4.60×10^{-28}	-0.02
2-623863-GTTC-G	rs60383093	*THEM18*	Body-mass index inverse-rank normalized	1.73×10^{-101}	0.06
20-63690302-G-A	rs2236506	*RTEL1-TNFRSF6B*	Atopic dermatitis	1.00×10^{-28}	0.10
3-45922676-T-C	rs11130078	*FYCO1*	Monocyte percentage of white cells	1.20×10^{-25}	0.02
3-71433933-G-A	rs62246017	*FOXP1*	Malignant neoplasm of skin excluding all cancers (controls excluding all cancers)	2.00×10^{-31}	0.08
5-132660808-A-C	rs848	*IL13*	Eosinophil counts	7.48×10^{-126}	-0.06
5-53277913-T-A	rs37807	*CTD-2143A15.1*	Corneal resistance factor (left)	5.67×10^{-22}	-0.14
6-32706117-C-T	rs1794269	*MTCO3P1*	Type 1 diabetes definitions combined	10.00×10^{-656}	0.87
8-142762416-C-T	rs10094888	*SLURP2. LYNX1*	Total testosterone levels	1.40×10^{-16}	-0.02
12-111446804-T-C	rs3184504	*SH2B3*	Eosinophil counts	2.23×10^{-308}	-0.10
12-122700003-T-C	rs612057	*RP11-324E6.6*	Plateletcrit	2.30×10^{-11}	0.01
12-52695529-G-A	rs1829637	*KRT77*	Basal cell carcinoma excluding all cancers (controls excluding all cancers)	2.51×10^{-08}	-0.04
13-53167993-A-G	rs1885767	*LINC01065*	Sleeplessness / insomnia	2.00×10^{-12}	-0.01
14-74711990-C-T	rs12434646	*AREL1*	Appendicular lean mass	8.42×10^{-18}	-0.02
16-11121178-A-G	rs34306440	*CLEC16A*	Eosinophil counts	1.62×10^{-105}	-0.05
16-31080754-G-A	rs7196726	*ZNF646*	Serum 25-Hydroxyvitamin D levels	2.23×10^{-308}	-0.01

16-53767042-T-C	rs1421085	*FTO*	Body-mass index inverse-rank normalized	$3.48x10^{-319}$	0.08
17-45617831-C-T	rs62066119	*LINC02210-CRHR1*	Mean spheric corpuscular volume	$9.00x10^{-128}$	-0.06
18-31339948-A-G	rs61730306	*DSG1*	Serum 25-Hydroxyvitamin D levels	$2.23x10^{-308}$	-0.02
19-10780563-C-A	rs2043332	*DNM2*	Statin medication	$1.94x10^{-47}$	-0.08
19-18015397-C-A	rs438735	*ARRDC2*	Monocyte count	$3.30x10^{-19}$	-0.03
20-63690302-G-A	rs2236506	*RTEL1-TNFRSF6B*	Atopic dermatitis	$1.00x10^{-28}$	0.10

Stratified LDSC suggests association to skin and immune tissues

Next, we performed a tissue-stratified linkage disequilibrium score regression (s-LDSC) analysis to identify which tissues are the most relevant for dermatophytosis infection (Figure 4, Table S5). We show top associations with connective tissue ($p = 4.32x10^{-10}$) and immune cells ($p = 1.96x10^{-8}$). As skin tissue is included in the connective tissue, the top associations strengthen our other results regarding the relevance of skin integrity and well-being and the internal immune defense defending against dermatophytosis fungal infections.

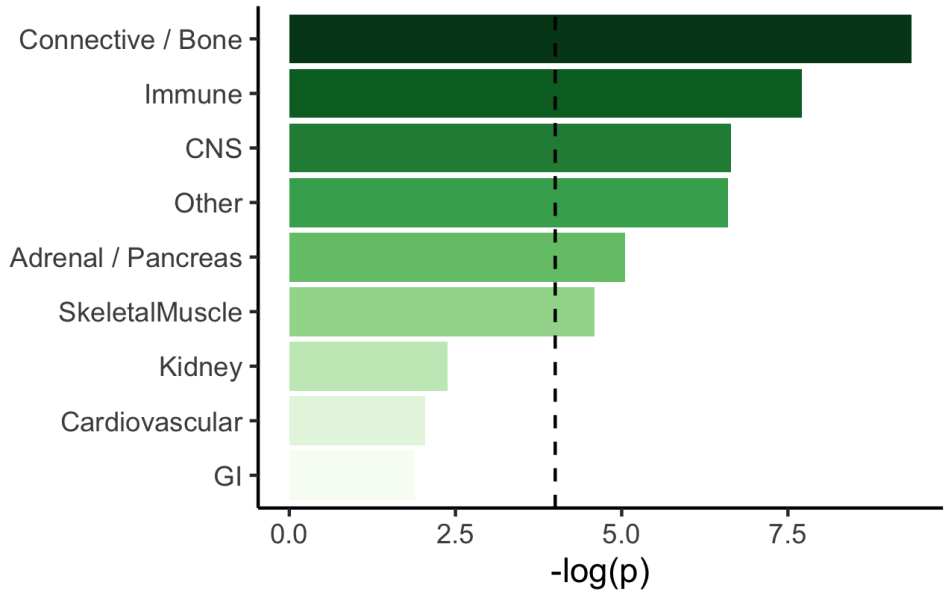

Figure 4. Stratified LDSC identifies connective tissue and immune cells as the most relevant tissue types for dermatophytosis infection.

To further study the specific related tissue types we performed s-LDSC multi-tissue analysis with 80 immune cell type and chromatin marker combinations and 39 skin cell and chromatin marker combinations [27]. We identified strongest association in the immune subset with T-helper cells (Primary T helper 17 cells PMA-I stimulated, H3K4me1, p = 0.0005) and in the skin subset with fibroblasts (Foreskin Fibroblast Primary Cells skin02, H3K27ac, p = 0.005) and keratinocytes (Foreskin

Keratinocyte Primary Cells skin03, H3K4me3, p = 0.024). Only the association with PMA-I stimulated primary T helper cells (H3K4me1) passes the multiple hypothesis corrected (Bonferroni-corrected) p-value threshold (threshold p-value 0.0006 for immune cells and 0.001 for skin cells) (Table S5).

Dermatophytosis shows genetic correlation with other diseases of skin and high BMI

We employed LDSC to study the genetic correlation of dermatophytosis with other infections of the skin and subcutaneous tissue, certain infectious and parasitic diseases and obesity. By using our meta-analysis summary statistics and all relevant endpoints in FinnGen we found associations with several skin disease endpoints (Figure 5, Table S7), many of which include itchiness, rash or redness of the skin. The most significant associations were with erysipelas (ICD10: A46, $p = 3.5 \times 10^{-20}$), other disorders of skin and subcutaneous tissue (ICD10: L98, $p = 7.4 \times 10^{-16}$) and infections of skin and subcutaneous tissue (ICD10: L00-L08, $p = 1.5 \times 10^{-10}$). Moreover, we observed shared genetic architecture between Scabies, Herpes simplex infection and other Bacterial diseases endpoints ($P < 0.0001$). The finding highlights the similar etiology behind skin-related diseases and proposes that the same genetic variants may increase our susceptibility to several types of skin diseases.

In addition, we studied the genetic correlation between dermatophytosis and BMI due to the association of several lead variants with changes in BMI and found a strong correlation with obesity ($p = 4.4 \times 10^{-49}$) and BMI IRN (inverse rank normalized) ($p = 3.2 \times 10^{-35}$).

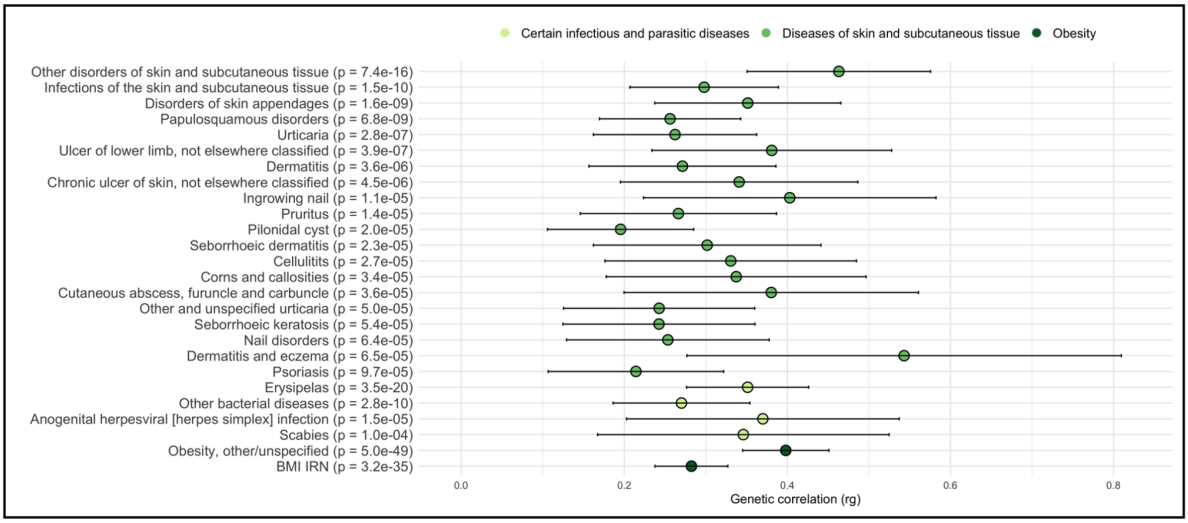

Figure 5. Statistically significant ($p < 0.05$) genetic correlations between dermatophytosis infection and FinnGen endpoints of other infections of skin and subcutaneous tissue, certain infectious and parasitic diseases and obesity. Lines represent 95% confidence intervals of the genetic correlation (rg).

Discussion

Our meta-analysis of over 250,000 cases and 1,300,000 controls identified 30 genetic loci associated with dermatophytosis, nine of which were either missense variants or in high LD with a missense variant. Several of the associated loci were linked to keratin processing and skin integrity, immune defense against pathogens, or environmental factors such as high BMI. Additionally, we identified skin and immune cells as the most relevant tissue types for dermatophytosis infection using stratified LDSC analysis for different tissue types (narrow tissue dataset). Finally, a genetic correlation analysis with other skin and subcutaneous tissue infections, parasitic diseases, and obesity revealed shared genetic architecture, highlighting commonalities between dermatophytosis, other skin diseases and high BMI. Overall, these findings implicate the role of immune mechanisms, environmental contributions from high BMI and vitamin D biology and most notably skin integrity and its barrier role, keratin biology and keratin processing in dermatophytosis.

Our findings suggest that genetic variation in keratin-related genes, including those influencing keratinocyte function and skin well-being—such as *KRT1*, *KRT77*, *FLG*, *SLURP2*, and *ZNF646*—play a critical role in susceptibility to dermatophyte infections. Since dermatophytes specifically target keratin using it as an energy source, it is noteworthy that genetic variation in these loci potentially can modulate disease susceptibility [19]. The findings highlight the role of keratin and barrier organs like the skin as the first line of defense against pathogens such as fungi.

We identified associations between dermatophytosis and genetic variation in several key keratin-related genes, including *KRT1*, *KRT77*, and *FLG*. *KRT1* encodes for Keratin 1 protein which is essential in maintaining the structural integrity of the skin by forming the cytoskeleton of keratinocytes, contributing to the skin's protective barrier against pathogens and environmental damage [28]. Moreover, Keratin 1 is part of the nutrient source for the fungal dermatophytes. Missense variation in *KRT1* may have two mechanisms that affect disease susceptibility. It is possible that the protective missense variant may influence the structure and stability of the keratin filaments improving the barrier function. Alternatively, the missense variant can affect the protein structure making it harder to be degraded by the fungi.

FLG is also directly involved in keratin processing, as it participates in the terminal differentiation of keratinocytes, where it aggregates keratin filaments into dense bundles forming a protective layer of dead keratinocytes on the outer surface of the skin. This layer, called stratum corneum, provides a tough and protective barrier both against mechanical stress and pathogens [29]. Moreover, filaggrin (*FLG*) is crucial in skin hydration, as its breakdown products serve as natural moisturizing factors (NMFs), retaining moisture in the skin and preventing dryness [30]. Additionally, these breakdown products help maintain a slightly acidic skin pH, inhibiting the growth of pathogenic microbiota [31]. Mutations in *KRT1* and *FLG* have been linked to skin disorders such as epidermolytic hyperkeratosis [32], atopic dermatitis [30], and ichthyosis vulgaris [31] .

Furthermore, colocalization analysis suggests the same genetic variant behind dermatophytosis and differential expression of *SLURP2* and *LYNX1* in the skin tissue. Both genes are expressed by keratinocytes and may play roles in skin homeostasis and the pathogenesis of skin disorders by

modulating nicotinic acetylcholine receptor functions [14]. Notably, *SLURP2* has been shown to promote keratinocyte hyperproliferation by inhibiting apoptosis. It also connects with the adaptive immune system and affects T-cell differentiation and activation. [33] Similar to *KRT1* and *FLG*, *SLURP2* has been linked to skin diseases, particularly Mal de Meleda, a rare genetic disorder causing thickened skin on the palms and soles [34].

Lastly, linked to skin well-being, our strongest association with dermatophytosis is with a missense variant in the *ZNF646* gene. *ZNF646* belongs to a family of zinc finger proteins that act in transcriptional regulation and protein degradation. This protein family is implicated in tissue development, particularly in the skin where several family members can modulate keratinocyte gene expression and differentiation [21]. Our findings suggest a regulatory effect of the *ZNF646* missense variant on *FLG* expression but how ZNF646 participates in dermatophytosis needs to be evaluated in future functional studies.

In addition to identifying genes linked directly to keratin processing and skin health, we showed a second group of lead variants associated with immune defense. The strongest association was found within the HLA region, which plays a critical role in defending against infections by presenting pathogen antigens to T-cells [23]. This finding highlights the involvement of canonical immune mechanisms in dermatophytosis.

Other notable immune-related associations were with *IL13* and *SH2B3*. Interleukin-13 (*IL13*) is involved in allergy, inflammation, and immune defense against parasites, such as helminths [24]. Interestingly, *IL13* also influences skin integrity by downregulating filaggrin and involucrin—two key components of the skin barrier. It also drives the release of itch-inducing molecules (pruritogens) and inflammatory cytokines, exacerbating chronic itch, a common symptom of dermatophytosis. *IL13* has been previously linked to atopic dermatitis [35]. *SH2B3* (*LNK*), on the other hand, plays a central role in modulating immune cell signaling by interacting with cytokine pathways and T/B cell receptor signaling, with associations to multiple autoimmune diseases [25].

A third group of genes identified in dermatophytosis GWAS were related to obesity, with the most notable being *FTO* and *IRX3*. *FTO* is a well-known gene responsible for energy homeostasis and body weight and mutations in its non-coding sequences are associated with obesity [15]. The functional effects of *FTO* variants are mediated through homeobox gene *IRX3* for which *FTO* acts as a long-range enhancer [15]. Additionally, variant annotations relate our missense variant in *ZNF646* with high BMI and obesity, further supporting the connection between metabolic health and susceptibility to dermatophytosis.

Our findings underscore the critical role of keratin processing and skin well-being, immune defense, and BMI-related genes in dermatophytosis susceptibility. Variant annotations further support these three categories. Tissue-type genetic correlation analysis identified immune cells and connective tissue as the most relevant tissue types in dermatophytosis infection. We also demonstrated significant genetic correlations between dermatophytosis and several other conditions, including skin and subcutaneous infections (such as other disorders of the skin and subcutaneous tissue, psoriasis and urticaria), certain infectious and parasitic diseases (erysipelas), and obesity (BMI).

Despite the substantial scale of our GWAS on dermatophytosis, several factors should be considered when interpreting our findings. Given that multiple skin infections can present with similar symptoms it is possible that we also capture variants that are not specific to dermatophytosis but implicate shared or overlapping biology with other diseases that share similar symptomatology. Second, although we increased sample size and power by utilizing data from multiple cohorts, our cases and controls are predominantly of European ancestry. While MVP and pan-UKB dataset includes individuals from other ancestries it is likely that we do not capture the full genetic architecture of dermatophytosis. Lastly, we observed that some lead variants were not consistently present across all four datasets, likely due to differences in genotyping arrays and imputation platforms, which may reduce the power to detect significant associations.

This study highlights the complex interplay between genetic factors involved in keratin production, immune response, and metabolic regulation in determining susceptibility to dermatophytosis. By identifying key genetic variants across multiple biobanks, we provide insights into the biological mechanisms underlying dermatophyte infection and shed light on potential targets for prevention and treatment. Our results underscore the importance of skin barrier integrity, immune defenses, and metabolic health in protecting against fungal infections. These findings pave the way for future research into precision medicine approaches for dermatophytosis, aiming to develop personalized strategies for individuals at higher genetic risk of infection.

Materials and methods

Cohorts

FinnGen is a population-based public-private population cohort established in 2017 [36]. The study combines genetic data with electronic health record data, including International Classification of Diseases (ICD) codes spanning an individual's entire lifespan, derived from primary care registers, hospital inpatient and outpatient visits and drug prescriptions of 520,000 participants. The project aims to improve understanding of the genetic etiology of diseases and disorders potentially leading to drug development.

The UK Biobank (UKB) is a prospective open-access study containing over 500,000 individuals aged 40-69 years upon entry to the study between 2006-1010 [37]. At the time of the entry to the cohort, a variety of health and lifestyle measures were collected, and blood and urine samples were taken for genetic and biochemistry analysis. Hospital in-patient (HES; N~470,000) and primary care (GP; N~231,000) records were later linked up to provide longitudinal data on disease diagnosis, operations, deaths, medications and deaths. In this study we used publicly available summary statistics from pan-UKBB study [38].

The Estonian Biobank is a population-based biobank with 212,955 participants [39]. Information on ICD-10 codes is obtained through regular linking with the National Health Insurance Fund and other relevant databases. The majority of the electronic health records have been collected since 2004.

The Million Veteran Project (MVP) is a longitudinal cohort study of diverse U.S. Veterans looking at how genes, lifestyle, military experiences, and exposures affect health and wellness [40]. It combines genetic data with electronic health records of 635,969 participants (data freeze 4) across four ethnic groups. In this study we extracted data from the publicly released summary statistics for MVP for Dermatophytosis for all ancestries (AFR, AMR, EAS, and EUR).

Ethics statements

Study subjects in FinnGen provided informed consent for biobank research, based on the Finnish Biobank Act. Alternatively, separate research cohorts, collected prior the Finnish Biobank Act came into effect (in September 2013) and start of FinnGen (August 2017), were collected based on study-specific consents and later transferred to the Finnish biobanks after approval by Fimea (Finnish Medicines Agency), the National Supervisory Authority for Welfare and Health. Recruitment protocols followed the biobank protocols approved by Fimea. The Coordinating Ethics Committee of the Hospital District of Helsinki and Uusimaa (HUS) statement number for the FinnGen study is Nr HUS/990/2017.

The FinnGen study is approved by Finnish Institute for Health and Welfare (permit numbers: THL/2031/6.02.00/2017, THL/1101/5.05.00/2017, THL/341/6.02.00/2018, THL/2222/6.02.00/2018, THL/283/6.02.00/2019, THL/1721/5.05.00/2019 and THL/1524/5.05.00/2020), Digital and population data service agency (permit numbers: VRK43431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3), the Social Insurance Institution (permit numbers: KELA 58/522/2017, KELA 131/522/2018, KELA 70/522/2019, KELA 98/522/2019, KELA 134/522/2019, KELA 138/522/2019, KELA 2/522/2020, KELA 16/522/2020), Findata permit numbers THL/2364/14.02/2020, THL/4055/14.06.00/2020, THL/3433/14.06.00/2020, THL/4432/14.06/2020, THL/5189/14.06/2020, THL/5894/14.06.00/2020, THL/6619/14.06.00/2020, THL/209/14.06.00/2021, THL/688/14.06.00/2021, THL/1284/14.06.00/2021, THL/1965/14.06.00/2021, THL/5546/14.02.00/2020, THL/2658/14.06.00/2021, THL/4235/14.06.00/2021, Statistics Finland (permit numbers: TK-53-1041-17 and TK/143/07.03.00/2020 (earlier TK-53-90-20) TK/1735/07.03.00/2021, TK/3112/07.03.00/2021) and Finnish Registry for Kidney Diseases permission/extract from the meeting minutes on 4th July 2019.

The Biobank Access Decisions for FinnGen samples and data utilized in FinnGen Data Freeze 11 include: THL Biobank BB2017_55, BB2017_111, BB2018_19, BB_2018_34, BB_2018_67, BB2018_71, BB2019_7, BB2019_8, BB2019_26, BB2020_1, BB2021_65, Finnish Red Cross Blood Service Biobank 7.12.2017, Helsinki Biobank HUS/359/2017, HUS/248/2020, HUS/430/2021 §28, §29, HUS/150/2022 §12, §13, §14, §15, §16, §17, §18, §23, §58, §59, HUS/128/2023 §18, Auria Biobank AB17-5154 and amendment #1 (August 17 2020) and amendments BB_2021-0140, BB_2021-0156 (August 26 2021, Feb 2 2022), BB_2021-0169, BB_2021-0179, BB_2021-0161, AB20-5926 and amendment #1 (April 23 2020) and it´s modifications (Sep 22 2021), BB_2022-0262, BB_2022-0256, Biobank Borealis of Northern Finland_2017_1013, 2021_5010, 2021_5010 Amendment, 2021_5018, 2021_5018 Amendment, 2021_5015, 2021_5015 Amendment, 2021_5015 Amendment_2, 2021_5023, 2021_5023 Amendment, 2021_5023 Amendment_2, 2021_5017, 2021_5017 Amendment, 2022_6001, 2022_6001 Amendment, 2022_6006 Amendment, 2022_6006 Amendment, 2022_6006

Amendment_2, BB22-0067, 2022_0262, 2022_0262 Amendment, Biobank of Eastern Finland 1186/2018 and amendment 22§/2020, 53§/2021, 13§/2022, 14§/2022, 15§/2022, 27§/2022, 28§/2022, 29§/2022, 33§/2022, 35§/2022, 36§/2022, 37§/2022, 39§/2022, 7§/2023, 32§/2023, 33§/2023, 34§/2023, 35§/2023, 36§/2023, 37§/2023, 38§/2023, 39§/2023, 40§/2023, 41§/2023, Finnish Clinical Biobank Tampere MH0004 and amendments (21.02.2020 & 06.10.2020), BB2021-0140 8§/2021, 9§/2021, §9/2022, §10/2022, §12/2022, 13§/2022, §20/2022, §21/2022, §22/2022, §23/2022, 28§/2022, 29§/2022, 30§/2022, 31§/2022, 32§/2022, 38§/2022, 40§/2022, 42§/2022, 1§/2023, Central Finland Biobank 1-2017, BB_2021-0161, BB_2021-0169, BB_2021-0179, BB_2021-0170, BB_2022-0256, BB_2022-0262, BB22-0067, Decision allowing to continue data processing until 31st Aug 2024 for projects: BB_2021-0179, BB22-0067,BB_2022-0262, BB_2021-0170, BB_2021-0164, BB_2021-0161, and BB_2021-0169, and Terveystalo Biobank STB 2018001 and amendment 25th Aug 2020, Finnish Hematological Registry and Clinical Biobank decision 18th June 2021, Arctic biobank P0844: ARC_2021_1001.

The activities of the EstBB are regulated by the Human Genes Research Act, which was adopted in 2000 specifically for the operations of the EstBB. Individual level data analysis in the EstBB was carried out under ethical approval 1.1-12/624 from the Estonian Committee on Bioethics and Human Research (Estonian Ministry of Social Affairs), using data according to release application 6-7/GI/33543 from the Estonian Biobank.

Genotyping and quality control

FinnGen R12 contains genetic data for 520 210 individuals. The samples were genotyped using Illumina (Illumina) and Affymetrix arrays (Thermo Fisher Scientific). The array consisted of 735,145 probes looking for 655,973 variants consisting of core backbone variants for imputation, rare coding variants enriched in the Finnish population, variants for KIR and HLA haplotypes, disease-specific markers and pharmacogenomic markers.

Genotyping data produced with previous chip platforms and reference genome builds were lifted over to build v.38 (GRCh38/hg38). For sample-wise quality control individuals exhibiting a discrepancy between genetically inferred sex and reported sex in registries, high genotype missingness (>5%), and excess heterozygosity (±4 standard deviations) were excluded. For variant-level QC, variants with high missingness (>2%), low Hardy–Weinberg equilibrium ($P < 1 \times 10–6$), and a minor allele count < 3 were filtered out. Chip-genotyped samples were pre-phased with Eagle 2.3.5 and imputed with the Finnish-specific SISu v4 imputation reference panel. Post-imputation quality control involved excluding variants with INFO score < 0.7 [36].

All EstBB participants have been genotyped at the Core Genotyping Lab of the Institute of Genomics, University of Tartu, using Illumina Global Screening Array v3.0_EST. Samples were genotyped and PLINK format files were created using Illumina GenomeStudio v2.0.4. Individuals were excluded from the analysis if their call rate was < 95%, if they were outliers of the absolute value of heterozygosity (> 3SD from the mean) or if sex defined based on heterozygosity of X chromosome did not match sex in phenotype data [41]. Before imputation, variants were filtered by call rate < 95%, HWE p-value < $1x10^{-4}$ (autosomal variants only), and minor allele frequency < 1%. Genotyped variant positions were in build 37 and were lifted over to build 38 using Picard. Phasing was performed using the Beagle v5.4 software

[42]. Imputation was performed with Beagle v5.4 software (beagle.22Jul22.46e.jar) and default settings. The dataset was split into batches of 5,000. A population-specific reference panel consisting of 2,695 WGS samples was utilized for imputation and standard Beagle hg38 recombination maps were used. Based on the principal component analysis, samples which were not from European ancestry individuals were removed. Duplicate and monozygous twin detection was performed with KING 2.2.7 [43], and one sample was removed from the pair of duplicates.

GWAS

GWAS in FinnGen was conducted using the REGENIE pipeline (https://github.com/FINNGEN/regenie-pipelines) adjusting for age, sex, chip, batch and ten first principal components. [44]

Association analysis in the Estonian Biobank was carried out for all variants with an INFO score > 0.4 using the additive model as implemented in REGENIE v3.0.3 with standard binary trait settings [44]. Logistic regression was carried out with adjustment for current age, age², sex and 10 PCs as covariates, analyzing only variants with a minimum minor allele count of 2.

UK biobank summary statistics were readily available from pan-UKBB study [38] and MVP summary statistics from their web server (https://phenomics.va.ornl.gov/pheweb/gia/meta/pheno/Phe_110).

The case and control definitions, numbers of cases and controls, used GWAS software and used covariates in each GWAS analysis are presented in Table 4.

Table 4. Cohorts, case and control definitions, numbers of cases and controls, used GWAS software and used covariates in each GWAS analysis included in the meta-analysis.

Cohort	Case definition	Control exclusions	N (cases)	N(controls	GWAS software	Covariates used
FinnGen	ICD10: B35	ICD10: B35	27662	471729	Regenie [44]	Age, sex, first 10 PCs, genotyping batches
UK biobank	ICD10: B35 ICD9: 110	ICD10: B35|B36|B37|B38|B39|B40|B41|B42|B43|B44|B45|B46|B47|B48|B49; ICD9: 110|111|112|113|114|115|116|117|118|119	27755	380368	Saige [45]	Age, sex, age * sex, age^2, age^2 * sex, first 10 PCs
Estonian biobank	ICD10: B35	ICD10: B35, B36, B37, B38, B39, B4	50241	106586	Regenie [44]	Gender, age, first 10 PCA and genotype batch control
MVP biobank	ICD9: 110	ICD9: 110	151164	413818	Saige [45]	Age, sex, and first 10 principal

						components

Meta-analysis

We conducted a meta-analysis with summary statistics from FinnGen, Estonian Biobank, UKB and MVP using the standard error method in METAL software [46]. The variants are annotated according to genome build 38.

The Manhattan plots for meta-analyses were plotted using R version 4.3.1 (packages: qqman and RColorBrewer).

Colocalization analysis

To assess the shared association of our lead variants to dermatophytosis and tissue-specific eQTLs, we performed colocalization analysis. For the analysis we used meta analysis summary statistics from dermatophytosis from a region +-50,000 base pairs around our lead variants and imported eQTL association statistics from GTEx [13] (https://gtexportal.org/home/) for the same region for all tissues.

Colocalization was performed using the R package coloc (v5.1.0.1 in R v4.2.2) [47] and co-localization plots were generated with LocusCompareR R package (v1.0.0) [48] using LD r^2 from 1000 Genomes [49] European-ancestry samples.

Tissue and cell type-specific analyses

To study relevant tissue and cell types for dermatophytosis infection, we employed stratified LDSC method [27]. First, we assessed relevant tissue types using cell-type groups data as used in Finucane et al. consisting of two gene expression datasets, GTEx project and 'Franke lab' dataset with 205 tissues and cell types, that were classified into nine categories (Connective/Bone, Immune, Other, CNS, Skeletal Muscle, Liver, Adrenal/Pancreas, Kidney, GI, Cardiovascular) [50-52]. Second, we studied relevant cell types using Multi_tissue_chromatin_1000Gv3_ldscores dataset from Finucane et al. (2018) composed of chromatin data from Roadmap Epigenomics and ENCODE projects. This dataset contains 489 tissue-specific chromatin-based annotations from peaks for six epigenetic marks (H3K27ac, H3K4me1, H3K4me3, H3K9ac, H3K36me3 and DHS) [53,54]. We used only annotations related to skin and immune cells, resulting in 80 immune cell type and chromatin marker combinations and 39 skin cell and chromatin marker combinations.

Genetic correlation

We performed genetic correlation between dermatophytosis infection and FinnGen endpoints of other infections of skin and subcutaneous tissue, certain infectious and parasitic diseases and obesity using the LD score regression method [55]. HapMap 3 SNP list and European LD score files, which are provided with the software, were used in our LD Score regression analyses. For dermatophytosis, we used summary statistics from our meta-analysis, and for FinnGen we used all endpoints in categories AB1 and L12 as well as selected BMI related endpoints (BMI_IRN and E4_OBESITYNAS). Further information on these FinnGen endpoints can be found at https://risteys.finregistry.fi/. When assessing the relevance of the genetic correlation, we adjusted p-value with Bonferroni correction. The forest plot for genetic correlation was generated using the ggplot2 package in R (version 4.4.1).

Acknowledgements

We want to acknowledge the participants and investigators of the FinnGen study. The FinnGen project is funded by two grants from Business Finland (HUS 4685/31/2016 and UH 4386/31/2016) and the following industry partners: AbbVie Inc., AstraZeneca UK Ltd, Biogen MA Inc., Bristol Myers Squibb (and Celgene Corporation & Celgene International II Sàrl), Genentech Inc., Merck Sharp & Dohme LCC, Pfizer Inc., GlaxoSmithKline Intellectual Property Development Ltd., Sanofi US Services Inc., Maze Therapeutics Inc., Janssen Biotech Inc, Novartis Pharma AG, and Boehringer Ingelheim International GmbH. Following biobanks are acknowledged for delivering biobank samples to FinnGen: Auria Biobank (www.auria.fi/biopankki), THL Biobank (www.thl.fi/biobank), Helsinki Biobank (www.helsinginbiopankki.fi), Biobank Borealis of Northern Finland (https://www.ppshp.fi/Tutkimus-ja-opetus/Biopankki/Pages/Biobank-Borealis-briefly-in-English.aspx), Finnish Clinical Biobank Tampere (www.tays.fi/en-US/Research_and_development/Finnish_Clinical_Biobank_Tampere), Biobank of Eastern Finland (www.ita-suomenbiopankki.fi/en), Central Finland Biobank (www.ksshp.fi/fi-FI/Potilaalle/Biopankki), Finnish Red Cross Blood Service Biobank (www.veripalvelu.fi/verenluovutus/biopankkitoiminta), Terveystalo Biobank (www.terveystalo.com/fi/Yritystietoa/Terveystalo-Biopankki/Biopankki/) and Arctic Biobank (https://www.oulu.fi/en/university/faculties-and-units/faculty-medicine/northern-finland-birth-cohorts-and-arctic-biobank). All Finnish Biobanks are members of BBMRI.fi infrastructure (www.bbmri.fi). Finnish Biobank Cooperative -FINBB (https://finbb.fi/) is the coordinator of BBMRI-ERIC operations in Finland. The Finnish biobank data can be accessed through the Fingenious® services (https://site.fingenious.fi/en/) managed by FINBB.

Equally, we want to acknowledge the participants of the Estonian Biobank for their contributions. The Estonian Genome Center analyses were partially carried out in the High Performance Computing Center, University of Tartu.

Funding

H.H. received funding for this project from Finland's Doctoral Education Pilot project.

The work of the E.A. was funded by the European Union through Horizon Europe research and innovation programs under grants no. 101137201 and 101137154, and Estonian Research Council Grant PRG1291.

Conflict of interest statement

Authors declare no conflict of interest.

References

1. Celestrino, G.A., Verrinder Veasey, J., Benard, G. & Sousa, M.G.T. Host immune responses in dermatophytes infection. *Mycoses* **64**, 477-483 (2021).
2. Hayette, M.-P. & Sacheli, R. Dermatophytosis, trends in epidemiology and diagnostic approach. *Current Fungal Infection Reports* **9**, 164-179 (2015).
3. Rouzaud, C*., et al.* Severe dermatophytosis and acquired or innate immunodeficiency: a review. *Journal of Fungi* **2**, 4 (2015).
4. Sacheli, R*., et al.* Epidemiology of dermatophytes in Belgium: A 5 years' survey. *Mycopathologia* **186**, 399-409 (2021).
5. Nenoff, P*., et al.* Mycology–an update part 2: dermatomycoses: clinical picture and diagnostics. *JDDG: Journal der Deutschen Dermatologischen Gesellschaft* **12**, 749-777 (2014).
6. Farhi, D., Savary, J., Pansart, S. & Hesse, S. Prospective study of feet onychomycosis in France: prevalence, clinical aspect, impact and management in general practice. (2011).
7. Gürcan, S., Tikveşli, M., Eskiocak, M., Kilic, H. & Otkun, M. Investigation of the agents and risk factors of dermatophytosis: a hospital-based study. *Mikrobiyoloji Bulteni* **42**, 95-102 (2008).
8. Petrucelli, M.F*., et al.* Epidemiology and diagnostic perspectives of dermatophytoses. *Journal of Fungi* **6**, 310 (2020).
9. Degreef, H. Clinical forms of dermatophytosis (ringworm infection). *Mycopathologia* **166**, 257-265 (2008).
10. Jazdarehee, A., Malekafzali, L., Lee, J., Lewis, R. & Mukovozov, I. Transmission of onychomycosis and dermatophytosis between household members: a scoping review. *Journal of Fungi* **8**, 60 (2022).
11. Maurano, M.T*., et al.* Systematic localization of common disease-associated variation in regulatory DNA. *Science* **337**, 1190-1195 (2012).
12. Adzhubei, I.A*., et al.* A method and server for predicting damaging missense mutations. *Nat Methods* **7**, 248-249 (2010).
13. Consortium, G. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318-1330 (2020).
14. Moriwaki, Y., Takada, K., Tsuji, S., Kawashima, K. & Misawa, H. Transcriptional regulation of SLURP2, a psoriasis-associated gene, is under control of IL-22 in the skin: A special reference to the nested gene LYNX1. *International immunopharmacology* **29**, 71-75 (2015).
15. Claussnitzer, M*., et al.* FTO obesity variant circuitry and adipocyte browning in humans. *New England Journal of Medicine* **373**, 895-907 (2015).
16. Son, J.H*., et al.* Risk factors of dermatophytosis among Korean adults. *Scientific reports* **12**, 13444 (2022).
17. Gupta, C*., et al.* Review on host-pathogen interaction in dermatophyte infections. *Journal of Medical Mycology* **33**, 101331 (2023).

18. Gnat, S., Łagowski, D. & Nowakiewicz, A. Genetic predisposition and its heredity in the context of increased prevalence of dermatophytoses. *Mycopathologia* **186**, 163-176 (2021).
19. Mercer, D.K. & Stewart, C.S. Keratin hydrolysis by dermatophytes. *Medical mycology* **57**, 13-22 (2019).
20. Karsch, S., Büchau, F., Magin, T.M. & Janshoff, A. An intact keratin network is crucial for mechanical integrity and barrier function in keratinocyte cell sheets. *Cellular and Molecular Life Sciences* **77**, 4397-4411 (2020).
21. Cassandri, M.*, et al.* Zinc-finger proteins in health and disease. *Cell death discovery* **3**, 1-12 (2017).
22. Patel, S., Xi, Z.F., Seo, E.Y., McGaughey, D. & Segre, J.A. Klf4 and corticosteroids activate an overlapping set of transcriptional targets to accelerate in utero epidermal barrier acquisition. *Proceedings of the National Academy of Sciences* **103**, 18668-18673 (2006).
23. Klein, J. & Sato, A. The HLA system. *New England journal of medicine* **343**, 702-709 (2000).
24. Brombacher, F. The role of interleukin-13 in infectious diseases and allergy. *Bioessays* **22**, 646-656 (2000).
25. Devallière, J. & Charreau, B. The adaptor Lnk (SH2B3): an emerging regulator in vascular cells and a link between immune and inflammatory signaling. *Biochemical pharmacology* **82**, 1391-1402 (2011).
26. Bikle, D.D. Vitamin D and the skin: Physiology and pathophysiology. *Reviews in Endocrine and Metabolic Disorders* **13**, 3-19 (2012).
27. Finucane, H.K.*, et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet* **50**, 621-629 (2018).
28. Roth, W.*, et al.* Keratin 1 maintains skin integrity and participates in an inflammatory network in skin through interleukin-18. *Journal of cell science* **125**, 5269-5279 (2012).
29. Sandilands, A., Sutherland, C., Irvine, A.D. & McLean, W.I. Filaggrin in the frontline: role in skin barrier function and disease. *Journal of cell science* **122**, 1285 (2009).
30. O'Regan, G.M., Sandilands, A., McLean, W.I. & Irvine, A.D. Filaggrin in atopic dermatitis. *Journal of Allergy and Clinical Immunology* **122**, 689-693 (2008).
31. Brown, S.J. & McLean, W.I. One remarkable molecule: filaggrin. *Journal of Investigative Dermatology* **132**, 751-762 (2012).
32. Virtanen, M., Vahlquist, A., Smith, S.K., Gedde-Dahl Jr, T. & Bowden, P.E. Splice site and deletion mutations in keratin (KRT1 and KRT10) genes: unusual phenotypic alterations in Scandinavian patients with epidermolytic hyperkeratosis. *Journal of investigative dermatology* **121**, 1013-1020 (2003).
33. Tsuji, H.*, et al.* SLURP-2, a novel member of the human Ly-6 superfamily that is up-regulated in psoriasis vulgaris. *Genomics* **81**, 26-33 (2003).
34. Allan, C.M.*, et al.* Palmoplantar keratoderma in Slurp2-deficient mice. *Journal of Investigative Dermatology* **136**, 436-443 (2016).
35. Furue, M. Regulation of filaggrin, loricrin, and involucrin by IL-4, IL-13, IL-17A, IL-22, AHR, and NRF2: pathogenic implications in atopic dermatitis. *International journal of molecular sciences* **21**, 5382 (2020).
36. Kurki MI, K.J., Palta P, Sipilä TP, Kristiansson K, Donner KM, Reeve MP, Laivuori H, Aavikko M, Kaunisto MA, Loukola A, Lahtela E, Mattsson H, Laiho P, Della Briotta Parolo P, Lehisto AA, Kanai M, Mars N, Rämö J, Kiiskinen T, Heyne HO, 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 AG, Sun B, Foley CN, Aalto-Setälä K, Alasoo K, Arvas M, Auro K, Biswas S, Bizaki-Vallaskangas A, Carpen O, Chen CY, Dada OA, Ding Z, Ehm MG, Eklund K, Färkkilä M, Finucane H, Ganna A, Ghazal A, Graham RR, Green EM, Hakanen A, Hautalahti M, Hedman ÅK, Hiltunen M, Hinttala R, Hovatta I, Hu X, Huertas-Vazquez A, Huilaja L, Hunkapiller J, Jacob H, Jensen JN, Joensuu H, John S, Julkunen V, Jung M, Junttila J, Kaarniranta K, Kähönen M, Kajanne R, Kallio L, Kälviäinen R, Kaprio J; FinnGen; Kerimov N,

Kettunen J, Kilpeläinen E, Kilpi T, Klinger K, Kosma VM, 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 SV, Niemi MEK, Niemi M, Niiranen T, O Donnell CJ, Obeidat ME, Okafo G, Ollila HM, Palomäki A, Palotie T, Partanen J, Paul DS, Pelkonen M, Pendergrass RK, Petrovski S, Pitkäranta A, Platt A, Pulford D, Punkka E, Pussinen P, Raghavan N, Rahimov F, Rajpal D, Renaud NA, Riley-Gillis B, Rodosthenous R, Saarentaus E, Salminen A, Salminen E, Salomaa V, Schleutker J, Serpi R, Shen HY, Siegel R, Silander K, Siltanen S, Soini S, Soininen H, Sul JH, Tachmazidou I, Tasanen K, Tienari P, Toppila-Salmi S, Tukiainen T, Tuomi T, Turunen JA, Ulirsch JC, Vaura F, Virolainen P, Waring J, Waterworth D, Yang R, Nelis M, Reigo A, Metspalu A, Milani L, Esko T, Fox C, Havulinna AS, Perola M, Ripatti S, Jalanko A, Laitinen T, Mäkelä TP, Plenge R, McCarthy M, Runz H, Daly MJ, Palotie A. FinnGen provides genetic insights from a well-phenotyped isolated population - PubMed. *Nature* **613**(2023 Jan).

37. Bycroft, C*., et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203-209 (2018).
38. Karczewski, K.J*., et al.* Pan-UK Biobank GWAS improves discovery, analysis of genetic architecture, and resolution into ancestry-enriched effects. *MedRxiv*, 2024.2003. 2013.24303864 (2024).
39. Leitsalu, L*., et al.* Cohort profile: Estonian biobank of the Estonian genome center, university of Tartu. *International journal of epidemiology* **44**, 1137-1147 (2015).
40. Verma, A*., et al.* Diversity and scale: Genetic architecture of 2068 traits in the VA Million Veteran Program. *Science* **385**, eadj1182 (2024).
41. Mitt, M*., et al.* Improved imputation accuracy of rare and low-frequency variants using population-specific high-coverage WGS-based imputation reference panel. *Eur J Hum Genet* **25**, 869-876 (2017).
42. Browning, B.L., Tian, X., Zhou, Y. & Browning, S.R. Fast two-stage phasing of large-scale sequence data. *Am J Hum Genet* **108**, 1880-1890 (2021).
43. Manichaikul, A*., et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867-2873 (2010).
44. Mbatchou, J*., et al.* Computationally efficient whole-genome regression for quantitative and binary traits. *Nature genetics* **53**, 1097-1103 (2021).
45. Zhou, W*., et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat Genet* **50**, 1335-1341 (2018).
46. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-2191 (2010).
47. Wallace, C. A more accurate method for colocalisation analysis allowing for multiple causal variants. *PLoS Genet* **17**, e1009440 (2021).
48. Liu, B., Gloudemans, M.J., Rao, A.S., Ingelsson, E. & Montgomery, S.B. Abundant associations with gene expression complicate GWAS follow-up. *Nat Genet* **51**, 768-769 (2019).
49. Auton, A*., et al.* A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).
50. Consortium, G. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648-660 (2015).
51. Pers, T.H*., et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* **6**, 5890 (2015).
52. Fehrmann, R.S*., et al.* Gene expression analysis identifies global gene dosage sensitivity in cancer. *Nat Genet* **47**, 115-125 (2015).
53. Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).
54. Kundaje, A*., et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-330 (2015).

55. Bulik-Sullivan, B.K.*, et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-295 (2015).