HHS Public Access

Author manuscript

Mycoses. Author manuscript; available in PMC 2024 February 05.

Published in final edited form as:

Mycoses. 2023 December ; 66(12): 1050–1055. doi:10.1111/myc.13649.

Association of Kallikrein Related Peptidase 3 (KLK3) gene with dermatophytosis in the UK biobank cohort

Steven Lehrer[1], **Peter H. Rheinstein**[2]

[1]Department of Radiation Oncology, Icahn School of Medicine at Mount Sinai, New York City, New York, USA

[2]Severn Health Solutions, Severna Park, Maryland, USA

Abstract

Background: In a previous genome wide association study (GWAS) of UK Biobank (UKB) data, we identified one susceptibility locus, tubulointerstitial nephritis antigen (TINAG), with genome wide significance for dermatophytosis. We used genotype calls from file UKB22418. These data are derived directly from Affymetrix DNA microarrays but are missing many genotype calls. Using computationally efficient approaches, UKB has entered imputed genotypes into a second dataset, UKB22828, increasing the number of testable variants by over 100-fold to 96 million variants.

Methods: In the current study, we used UKB imputed genotypes in UKB22828 to identify dermatophytosis susceptibility loci. To identify cases of dermatophytosis, we used ICD10 code B35, which covers tinea barbae, tinea capitis, tinea unguium, tinea manuum, tinea pedis, tinea corporis, tinea imbricata, tinea cruris, other dermatophytoses and dermatophytosis, unspecified. We used PLINK, a whole-genome association analysis toolset, to analyse the UKB22828 chromosome files.

Results: GWAS summary (Manhattan) plot of the meta-analysis association statistics highlighted two susceptibility loci, TINAG and Kallikrein Related Peptidase 3 (KLK3), with genome wide significance for dermatophytosis. KLK3, also known as prostate specific antigen (PSA), belongs to a subclass of serine proteases with a variety of physiological functions.

Conclusion: KLK3 may be a dermatophytosis susceptibility gene. KLK3 could affect risk of dermatophytosis, since kallikreins are necessary for normal homeostasis of the skin.

Correspondence: Steven Lehrer, Box 1236 Radiation Oncology, Mount Sinai Medical Center, 1 Gustave L. Levy Place, New York 10029, USA. steven.lehrer@mssm.edu.

AUTHOR CONTRIBUTIONS

Steven Lehrer: Conceptualization. **Peter H. Rheinstein:** Conceptualization.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare.

ETHICS STATEMENT

UK Biobank has approval from the Northwest Multi-center Research Ethics Committee (MREC), which covers the UK. It also sought the approval in England and Wales from the Patient Information Advisory Group (PIAG) for gaining access to information that would allow it to invite people to participate. PIAG has since been replaced by the National Information Governance Board for Health & Social Care (NIGB). In Scotland, UK Biobank has approval from the Community Health Index Advisory Group (CHIAG).

Keywords

dermatophytosis; genes; GWAS; mycoses; polygenic

1 | INTRODUCTION

Individual factors, including genetics, predispose to dermatophytosis, a common condition. Acute inflammatory forms of dermatophytosis can have no symptoms at all or can be life-threatening. Due to the high cost of therapy, dermatophytosis is a substantial financial burden.[1,2]

A retrospective analysis of the hereditary propensity to dermatophytoses revealed at least several proven genetic relationships such as race, CARD9 deficiency, HLA-DR4 and HLA-DR8 type and genes encoding interleukin-22, β-defensin 2 and 4 as well as genetic defects in dectin-1, which increased the prevalence of dermatophytosis in families and were involved in the inheritance of susceptibility in family members.[1] Deficiency of SERPINB7 (Nagashima-type palmoplantar keratosis) and SERPINA12 both result in palmoplantar keratoderma with high susceptibility to plantar dermatophytosis.[3]

In a previous genome wide association study (GWAS) of UK Biobank (UKB) data, we identified one susceptibility locus, tubulointerstitial nephritis antigen (TINAG), with genome wide significance for dermatophytosis.[4] The top SNP was rs16885197, a missense variant within TINAG, position chr6:54308557, alleles A > G, minor allele frequency (MAF) 0.014. Multivariate logistic regression indicated that the minor G allele increased odds ratio of dermatophytosis by 7.8. Carrying two G alleles raised dermatophytosis odds ratio by a factor of 14.

We used genotype calls from file UKB22418. These data are derived directly from Affymetrix DNA microarrays but are missing many genotype calls.[5]

The estimation of missing genotype calls using statistical inference is known as genotype imputation. Imputation is increasingly enhancing the number of SNPs accessible in data, not just to fill in gaps left by genotyping errors but also to estimate the genotypes of variants that were not directly tested. Using computationally efficient approaches together with the Haplotype Reference Consortium and UK10K haplotype resources, UKB has entered imputed genotypes into a second dataset, UKB22828, increasing the number of testable variants by over 100-fold to 96 million variants.[6]

In the current study, we used UKB imputed genotypes in UKB22828 to identify dermatophytosis susceptibility loci.

2 | METHODS

The UK Biobank is a large prospective observational study of men and women. Participants were recruited from across 22 centers located throughout England, Wales and Scotland between 2006 and 2010 and continue to be longitudinally followed for capture of sub-

sequent health events.[7] Follow-up health information is provided by linkage to primary care electronic health records, death and cancer registries and hospital admission records.[8]

Our UK Biobank application was approved as UKB project 57245 (S.L., P.H.R.). To identify cases of dermatophytosis, we used ICD10 code B35, which covers tinea barbae, tinea capitis, tinea unguium, tinea manuum, tinea pedis, tinea corporis, tinea imbricata, tinea cruris, other dermatophytoses and dermatophytosis, unspecified.

Data processing was performed on Minerva, a Linux mainframe with Centos 7.6, at the Icahn School of Medicine at Mount Sinai. We used PLINK, a whole-genome association analysis toolset, to analyse the UKB22828 chromosome files[9] and the UK Biobank Data Parser (ukbb parser), a python-based package that allows easy interfacing with the large UK Biobank dataset.[10] We used the R package qqman for the Manhattan and qq plots.[11] Other statistical analyses were done with R and SPSS 26.

We followed quality control procedures[12] that consisted of the following:

1. Missingness of SNPS 0.05: This command excluded SNPs that are missing in a large proportion of the subjects. In this step, SNPs with low genotype calls were removed.
2. Missingness of individuals 0.05: This command excluded individuals who had high rates of genotype missingness. In this step, individuals with low genotype calls were removed.
3. Hardy Weinberg equilibrium 1e-6: This command excluded markers which deviate from Hardy–Weinberg equilibrium.
4. Minor allele frequency (MAF) threshold 0.01: This command included only SNPs above the set MAF threshold.

3 | RESULTS

We analysed data from 462,737 subjects. The age at enrollment was 56 ± 8 (mean ± SD). The subjects were 56% women, 44% men, 95% white and British, 15 ± 5 years of education. Five hundred and thirty-two subjects had dermatophytosis.

GWAS summary (Manhattan) plot of the meta-analysis association statistics, highlighting two susceptibility loci, Kallikrein Related Peptidase 3 (KLK3) and TINAG, with genome wide significance for dermatophytosis, is shown in Figure 1. The upper horizontal line indicates the genome wide significance threshold of a p value less than 5×10^{-8}. Figure 1 shows that Kallikrein Related Peptidase 3 (KLK3) and TINAG are the only susceptibility loci that met the p value threshold of less than 5×10^{-8}. rs61729813 is the KLK3 SNP that was found to be most significantly associated with dermatophytosis in GWAS.

The qq plot is shown in Figure 2. The x-axis represents expected −log10 (p), y-axis observed −log10 (p) of each SNP. The genomic inflation factor (lambda gc) =1.1. Values up to 1.10 are generally considered acceptable for GWAS and suggest no systematic biases.

Figure 3 shows the LocusZoom plot of KLK3 association. Genomic position is depicted on the x-axis. The left y-axis shows the −log10 of the *p*-value. SNPs are coloured based on their correlation (r^2) with the labelled top SNP rs61729813 (purple diamond), which has the smallest *p* value in the region. The fine-scale recombination rates estimated from 1000 Genomes (EUR) data (right y axis) are indicated by the fluctuating blue line. The position of KLK3 relative to rs61729813 is displayed. Figure 3B shows the LocusZoom plot of TINAG association. The labelled top SNP rs16885197 (purple diamond) has the smallest *p* value in the region. Figure 3 and 3B indicate that Kallikrein Related Peptidase 3 (KLK3) and TINAG are the only susceptibility loci that met the *p* value threshold of less than 5×10^{-8}.

In the GWAS rs61729813 was the SNP most significantly associated with dermatophytosis. rs61729813 is a missense variant within an exon of KLK3, alleles C > G, minor allele frequency (MAF) 0.012. Table 1 contains genotype of rs61729813 versus phenotype, unaffected or mycosis (dermatophytosis) in 462,737 subjects. 0.1% of subjects with genotype CC had dermatophytosis, 0.3% of subjects with genotype CG had dermatophytosis, no subjects were homozygous GG (Fisher exact test two tailed $p < 0.001$).

Results of logistic regression are in Table 2. The dermatophytosis odds ratio (O.R.) for males was 2.19, indicating dermatophytosis is more common in men. The O.R. for age of dermatophytosis was 1.034, in other words dermatophytosis incidence increases with every year of age. Diabetes type 2 increased risk of dermatophytosis, O.R. 2.532. Subjects that were carriers of the G allele of rs61729813 were at increased risk of dermatophytosis (O.R. 2.321).

4 | DISCUSSION

KLK3, also known as prostate specific antigen (PSA), belongs to a subclass of serine proteases with a variety of physiological functions. A growing body of research indicates that numerous kallikreins have a role in carcinogenesis and that some of them may serve as innovative cancer and other disease biomarkers. The KLK3 gene is one of 15 members of the kallikrein subfamily that are grouped together on chromosome 19. KLK3 encodes a protease, a single-chain glycoprotein, which is produced in the prostate gland's epithelial cells and found in seminal plasma. KLK3 protein is believed to hydrolyze high molecular mass seminal vesicle protein during the liquefaction of seminal coagulum. Serum PSA level aids in the detection and follow-up of prostatic cancer. KLK3 gene alternative splicing results in many transcript variants that each encode a different isoform.[13]

KLK3 could affect risk of dermatophytosis, since kallikreins are necessary for normal homeostasis of the skin. Keratinocytes of the upper stratum granulosum secrete kallikreins into the stratum corneum. In addition to their function in skin renewal, kallikreins control skin inflammation, the epidermal lipid-rich permeability barrier, and innate immune responses in the epidermis. Kallikreins degrade lipid-processing enzymes and activate the inflammatory response mediator PAR2 on the cell surface of keratinocytes. Antimicrobial cathelicidins, smaller antimicrobial peptides and pro-inflammatory cytokines all interact with kallikreins. In various skin conditions, including psoriasis, atopic dermatitis, acne rosacea and Netherton syndrome (NS), kallikrein regulatory function is disrupted.[14]

Dermatophytosis is more common in men and diabetics.[15,16] Our analysis confirms this association and indicates that it is independent of the significant influence of KLK3 (Table 2). Of interest is the fact that high dose itraconazole, an antifungal used to treat dermatophytosis, has modest antitumor activity in men with metastatic castration resistant prostate cancer that is not mediated by testosterone suppression. The effects of itraconazole appeared to be associated with inhibition of Hedgehog signalling in skin biopsies.[17]

Principal Component Analysis of UKBB data was published previously. The x and y axes of variation reduce the data to a small number of dimensions, describing as much variability as possible; they are defined as the top eigenvectors of a covariance matrix between samples.[18] The UKB subjects appear to cluster as a single group, indicating race is not a confounding variable in the GWAS, corroborating the lambda value above. Even though the UK Biobank cohort includes many participants from a variety of ethnic backgrounds, GWAS is possible without sacrificing adequate sample size because most UK Biobank cohort participants report their ethnic background as *British*, within the broader-level group *white*, 88%.[19]

Weaknesses in our study:

- We are uncertain why dermatophytosis incidence in the UK Biobank is so low (532/462737). Other sources rank its prevalence in the U.K. as high as 20%.[20,21] Perhaps most people with dermatophytosis do not bother to report it to a healthcare professional; or UK Biobank is only capturing patients who have severe cases.
- Heterogeneously localised cases are included such as onychomycosis and tinea capitis, although these clinical entities may be associated with different loci.
- UKBB data do not specify diagnostic criteria of dermatophytosis.
- KLK3 expression in skin is low.[22]
- The GWAS only suggests the association of the region 19q13.3 with dermatophytosis susceptibility. Our study does not prove that KLK3 is associated with susceptibility. Another KLK gene clustered at the same locus could be the susceptibility gene. The expression of the genes located around the locus should be evaluated in the skin, using ample public data, including single cell RNA analysis.
- It will be necessary to study and analyse the genotypes of dermatologist patients with clinically and laboratory confirmed dermatophytosis, especially in cases of recurrent familial dermatophytosis. The current study could serve as an important part or addition to the work already done.

In conclusion, KLK3 may be a dermatophytosis susceptibility gene. More research into genetic and other predisposing factors for dermatophytosis is critical because of the implications for prophylaxis and therapy. It might be possible to prevent infection and recurrence by identifying people who are vulnerable to chronic dermatophytosis. Identifying high-risk families would enable their members to be educated about the dangers of fungal diseases.

ACKNOWLEDGEMENTS

This work was supported in part through the computational resources and staff expertise provided by Scientific Computing at the Icahn School of Medicine at Mount Sinai. Research reported in this paper was also supported by the Office of Research Infrastructure of the National Institutes of Health under award numbers S10OD018522 and S10OD026880. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

DATA AVAILABILITY STATEMENT

Data in this study may be obtained from UK Biobank after approval of a formal application. See this link for additional information https://www.ukbiobank.ac.uk/. Posted as a preprint https://www.medrxiv.org/content/10.1101/2022.10.09.22280866v1.

REFERENCES

1. Gnat S, Łagowski D, Nowakiewicz A. Genetic predisposition and its heredity in the context of increased prevalence of Dermatophytoses. Mycopathologia. 2021;186:163–176. [PubMed: 33523393]
2. Son JH, Doh JY, Han K, et al. Risk factors of dermatophytosis among Korean adults. Sci Rep. 2022;12:13444. [PubMed: 35927302]
3. Harjama L, Karvonen V, Kettunen K, et al. Hereditary palmoplantar keratoderma—phenotypes and mutations in 64 patients. J Eur Acad Dermatol Venereol. 2021;35:1874–1880. [PubMed: 33914963]
4. Lehrer S, Rheinstein PH. Genome-wide association study of dermatophytosis in the UK biobank cohort. J Eur Acad Dermatol Venereol. 2022;36:2482–2487. [PubMed: 35796184]
5. Bycroft C, Freeman C, Petkova D, et al. Genome-wide genetic data oñ500,000 UK biobank participants. bioRxiv. 2017;166298.
6. Collister JA, Liu X, Clifton L. Calculating polygenic risk scores (PRS) in UK biobank: a practical guide for epidemiologists. Front Genet. 2022;13:818574. [PubMed: 35251129]
7. Arthur RS, Wang T, Xue X, Kamensky V, Rohan TE. Genetic factors, adherence to healthy lifestyle behavior, and risk of invasive breast cancer among women in the UK biobank. J Natl Cancer Inst. 2020;112:893–901. [PubMed: 31899501]
8. Khanji MY, Aung N, Chahal CAA, Petersen SE. COVID-19 and the UK biobank—opportunities and challenges for research and collaboration with other large population studies. Front Cardiovasc Med. 2020;7:156. [PubMed: 33195444]
9. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7. [PubMed: 25722852]
10. Zhu A, Salminen LE, Thompson PM, Jahanshad N. The UK biobank data parser: a tool with built in and customizable filters for brain studies. Organization for Hum Brain Mapp. 2019;9–13.
11. Turner SD. Qqman: an R package for visualizing GWAS results using Q-Q and Manhattan plots. bioRxiv. 2014;005165.
12. Marees AT, de Kluiver H, Stringer S, et al. A tutorial on conducting genome-wide association studies: quality control and statistical analysis. Int J Methods Psychiatr Res. 2018;27:e1608. [PubMed: 29484742]
13. Koistinen H, Künnapuu J, Jeltsch M. KLK3 in the regulation of angiogenesis—tumorigenic or not? Int J Mol Sci. 2021;22:13545. [PubMed: 34948344]
14. Avgeris M, Scorilas A. Kallikrein-related peptidases (KLKs) as emerging therapeutic targets: focus on prostate cancer and skin pathologies. Expert Opin Ther Targets. 2016;20:801–818. [PubMed: 26941073]
15. Oz Y, Qoraan I, Oz A, Balta I. Prevalence and epidemiology of tinea pedis and toenail onychomycosis and antifungal susceptibility of the causative agents in patients with type 2 diabetes in Turkey. Int J Dermatol. 2017;56:68–74. [PubMed: 27667305]
16. Sentamilselvi G, Kamalam A, Ajithadas K, Janaki C, Thambiah A. Scenario of chronic dermatophytosis: an Indian study. Mycopathologia. 1997;140:129–135. [PubMed: 9691500]

17. Antonarakis ES, Heath EI, Smith DC, et al. Repurposing itraconazole as a treatment for advanced prostate cancer: a noncomparative randomized phase II trial in men with metastatic castration-resistant prostate cancer. Oncologist. 2013;18:163–173. [PubMed: 23340005]
18. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006;38:904–909. [PubMed: 16862161]
19. Bycroft C, Freeman C, Petkova D, et al. The UK biobank resource with deep phenotyping and genomic data. Nature. 2018;562:203–209. [PubMed: 30305743]
20. Roberts D Prevalence of dermatophyte onychomycosis in the United Kingdom: results of an omnibus survey. Br J Dermatol. 1992;126:23–27. [PubMed: 1531924]
21. Havlickova B, Czaika VA, Friedrich M. Epidemiological trends in skin mycoses worldwide. Mycoses. 2008;51:2–15. [PubMed: 18783559]
22. Shaw JL, Diamandis EP. Distribution of 15 human kallikreins in tissues and biological fluids. Clin Chem. 2007;53:1423–1432. [PubMed: 17573418]

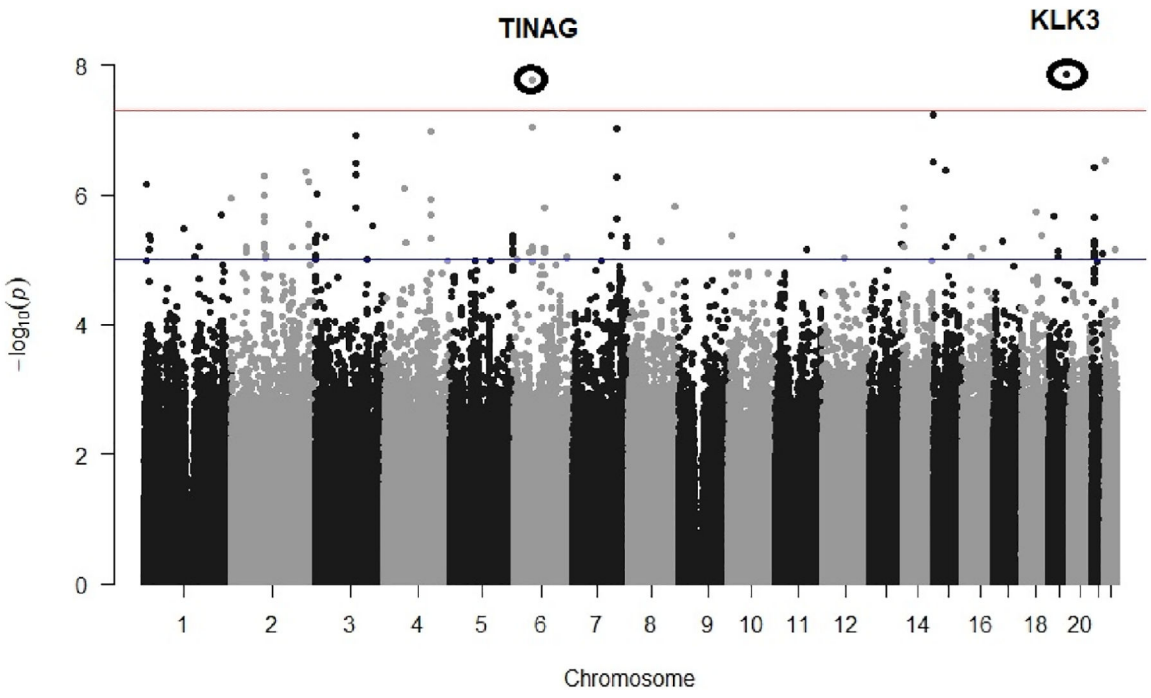

FIGURE 1.

Manhattan plot showing significantly associated dermatophytosis risk loci for TINAG, chromosome 6p12.1, and KLK3, chromosome 19q13.3.

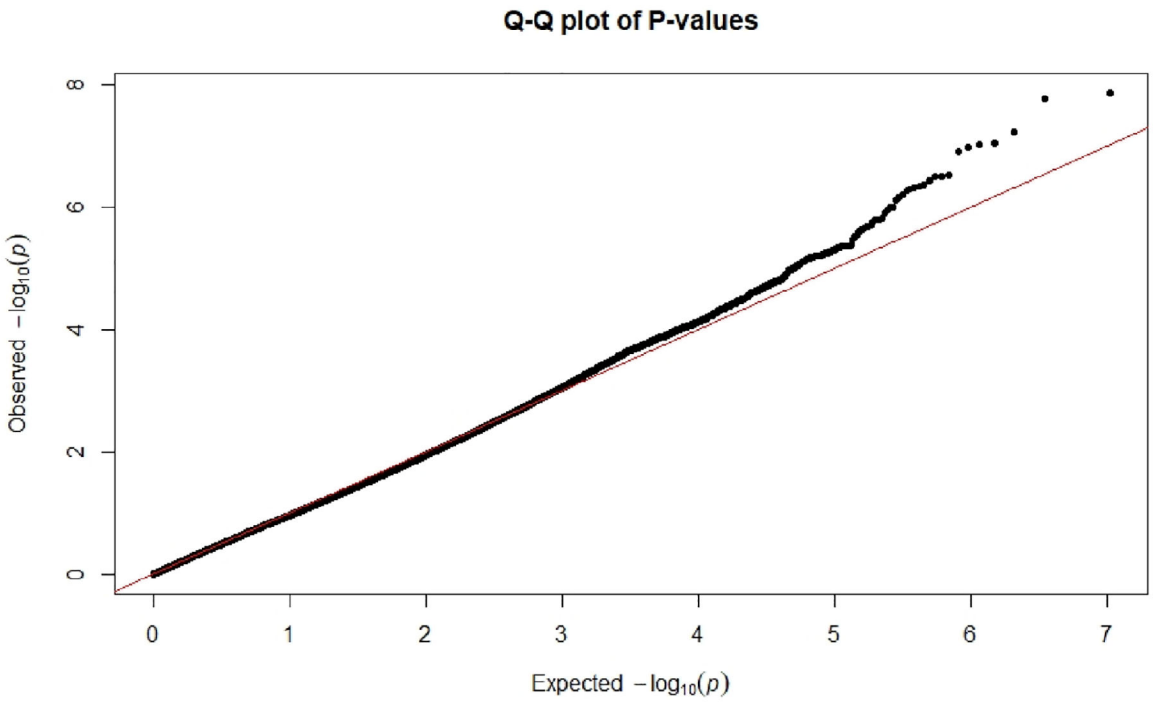

FIGURE 2.

qq plot of *p* values from GWAS data. Note that most of the *p*-values observed follow a uniform distribution (left segment of line) but the few that are in linkage disequilibrium with causal polymorphisms produce significant *p*-values (upper right segment of line). The genomic inflation factor (lambda gc) is 1.1. Values up to 1.1 are generally considered acceptable for GWAS and suggest no systematic biases.

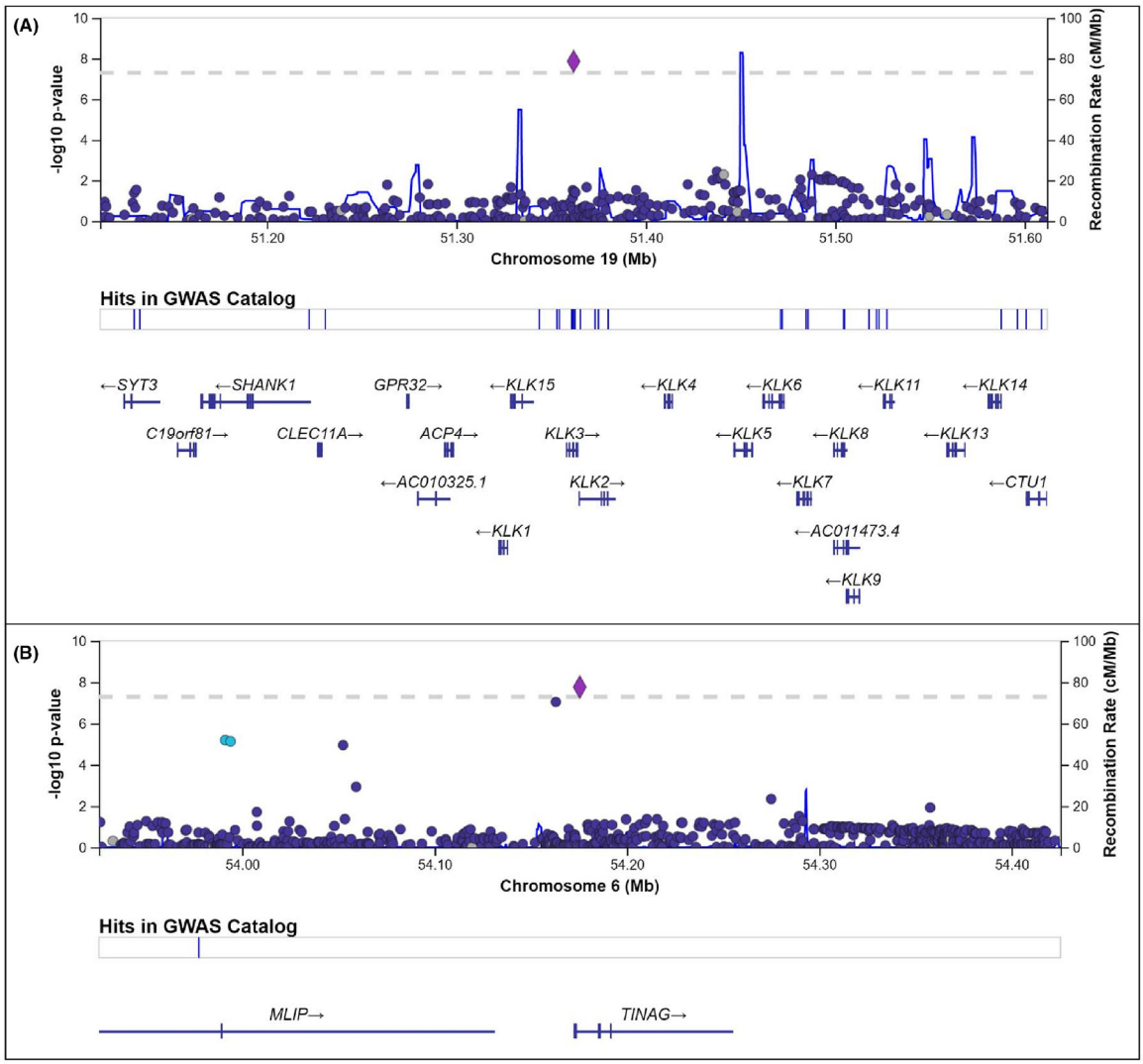

FIGURE 3.

(A)LocusZoom plot of KLK3 association. Genomic position is depicted on the x-axis. The left y-axis shows the $-\log_{10}$ of the *p*-value. SNPs are coloured based on their correlation (r^2) with the labelled top SNP rs61729813 (purple diamond), which has the smallest *p* value in the region. The fine-scale recombination rates estimated from 1000 Genomes (EUR) data (right y axis) are indicated by the fluctuating blue line. The position of KLK3 relative to rs61729813 is displayed. (B) LocusZoom plot of TINAG association. The labelled top SNP rs16885197 (purple diamond) has the smallest *p* value in the region.

TABLE 1

Genotype of rs61729813 versus phenotype, unaffected or mycosis (dermatophytosis) in 462,737 subjects. 0.1% of subjects with genotype CC had dermatophytosis, 0.3% of subjects with genotype CG had dermatophytosis, no subjects were homozygous GG. Fisher exact test two tailed $p < 0.001$.

Genotype		neg	Mycosis	Total
CC	N	451,111	503	451,614
	%	99.90%	0.10%	100.00%
CG	N	11,094	29	11,123
	%	99.70%	0.30%	100.00%
all	N	462,205	532	462,737
	%	99.90%	0.10%	100.00%

TABLE 2

Logistic regression with 95% confidence intervals, lower bound (L.B.), upper bound (U.B.). Independent variables sex, age, diabetes type 2, genotype CC versus CG; dermatophytosis present or absent, dependent variable. The dermatophytosis odds ratio (O.R.) for males was 2.19, indicating dermatophytosis is more common in men. The O.R. for age of dermatophytosis is 1.034, in other words dermatophytosis incidence increases with every year of age. Diabetes type 2 increased risk of dermatophytosis, O.R. 2.532. Subjects that were carriers of the G allele of rs61729813 were at increased risk of dermatophytosis (O.R. 2.321).

	95% L.B.	**O.R.**	**95% U.B.**	***p* value**
Sex	1.831	2.192	2.626	<.001
Age	1.021	1.034	1.046	<.001
Diabetes type 2	2.001	2.532	3.203	<.001
Genotype	1.596	2.321	3.375	<.001