Deep intronic 9q21.11 polymorphism contributes to atopic dermatitis risk through methylation regulated expression of tight junction protein 2

Running title: AD-associated variant alters epigenetics

Eliza Lim YY[1], Yie Sio Y[1], Say YH[1,2,3†], Reginald K[3], Tim Chew F[1*]

[1]Department of Biological Sciences, Faculty of Science, National University of Singapore, Singapore

[2]Department of Biomedical Science, Faculty of Science, Universiti Tunku Abdul Rahman (UTAR) Kampar Campus, Kampar, Perak, Malaysia

[3]Department of Biological Sciences, School of Medical and Life Sciences, Sunway University, Malaysia

[*]**Corresponding author:**
Fook Tim Chew
Associate Professor, Department of Biological Sciences,
Faculty of Science, National University of Singapore,
Allergy and Molecular Immunology Laboratory,
Lee Hiok Kwee Functional Genomics Laboratories,
Block S2, Level 5, 14 Science Drive 4,
Lower Kent Ridge Road, Singapore 117543.
E-mail: dbscft@nus.edu.sg

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.18176/jiaci.0978

Abstract

Background: Atopic dermatitis (AD) is a chronic inflammatory itchy skin condition. Genomic- and epigenetic wide association studies provide insights into the genetic susceptibility and potential underlying disease pathogenesis.

Objective: This study sought to functionally characterise an AD-associated single nucleotide polymorphism (SNP) located deep intronic of the *tight junction protein 2* (*TJP2*) gene (9q21.11 locus), identified through a genome-wide association study (GWAS).

Methods: The association between the 9q21.11 locus (rs7872806) and AD was identified through a GWAS of 956 cases and 723 controls. *TJP2* expression in peripheral blood mononuclear cells (PBMCs) was assessed against the rs7872806 genotypes. Allele-specific methylation was evaluated at CpG sites 10kb up- and down-stream of the 9q21.11 locus. Effects of DNA methylation on *TJP2* expression was validated via *in vitro* methylation and 5-aza-2'-deoxycytidine-induced transcriptional activation studies. Trans-epidermal water loss measurements were used to determine skin barrier function.

Results: The major allele of rs7872806 was determined to increase AD risk by 2.64-fold (adjusted p-value=2.40 x 10^{-18}, OR=0.38), associated with increased methylation levels at cg13920460 site (p<0.001) and lower *TJP2* expression in PBMCs (Pearson's p=1.09 x 10^{-6}, Pearson's R=-0.313, p<0.001). Methylation inhibition by 5-aza-2'-deoxycytidine increased *TJP2* promoter activity by up to 85%. Elimination of the cg13920460 methylation site increased expression by approximately 25%. The rs7872806 major allele was also found to be associated with increased trans-epidermal water loss (p<0.001).

Conclusion: Epigenetic influence at CpG site cg13920460 is associated with rs7872806 located deep intronic at 9q21.11. The SNP confers susceptibility to AD through altering *TJP2* expression and promoting trans-epidermal water loss.

Key words: Atopic dermatitis. Genome-wide association studies. Epigenetics. Methylation. Expression quantitative trait loci. Tight junction protein. Transepidermal water loss. Epidermal hyperplasia.

Resumen

Antecedentes: La dermatitis atópica (DA) es una enfermedad inflamatoria crónica de la piel que produce picor. Los estudios de asociación genómica y epigenética proporcionan información sobre la susceptibilidad genética y la posible patogénesis subyacente de la enfermedad.

Objetivo: El objetivo de este estudio es caracterizar funcionalmente un polimorfismo de nucleótido único (SNP) asociado a la DA y localizado en la parte intrónica profunda del gen de la proteína de unión estrecha 2 (TJP2) (locus 9q21.11), identificado mediante un estudio de asociación de genoma completo (GWAS).

Métodos: La asociación entre el locus 9q21.11 (rs7872806) y la DA se identificó mediante el estudio GWAS de 956 casos y 723 controles. Se evaluó la expresión de TJP2 en células mononucleares de sangre periférica (PBMCs) frente a los genotipos rs7872806. Se evaluó la metilación específica del alelo en sitios CpG 10kb corriente arriba y abajo del locus 9q21.11. Los efectos de la metilación del ADN en la expresión de TJP2 se validaron mediante estudios de metilación in vitro y de activación transcripcional inducida por 5-aza-2'-desoxicitidina. Para determinar la función de barrera del pie, se utilizaron mediciones de pérdida de agua transepidérmica

Resultados: Se determinó que el alelo principal del rs7872806 aumentaba el riesgo de DA 2,64 veces (p-valor ajustado= 2,40 x 10-18, OR= 0,38), asociado con mayores niveles de metilación en el sitio cg13920460 (p< 0,001) y menor expresión de TJP2 en PBMCs (p de Pearson= 1,09 x 10-6, R de Pearson= -0,313, p< 0,001). La inhibición de la metilación mediante 5-aza-2'-desoxicitidina aumentó la actividad promotora de TJP2 hasta en un 85%. La eliminación del sitio de metilación cg13920460 aumentó la expresión en aproximadamente un 25%. El alelo principal rs7872806 también se asoció con una mayor pérdida de agua transepidérmica (p< 0,001).

Conclusión: La influencia epigenética en el sitio CpG cg13920460 está asociada con rs7872806 localizado intrónicamente en 9q21.11. El SNP confiere susceptibilidad a la DA a través de la alteración de la expresión de TJP2 y la promoción de la pérdida de agua transepidérmica.

Palabras clave: Dermatitis atópica. Estudios de asociación del genoma completo. Epigenética. Metilación. Loci de rasgos cuantitativos de expresión. Proteína de unión estrecha. Pérdida transepidérmica de agua. Hiperplasia epidérmica.

doi: 10.18176/jiaci.0978

Summary Box

What do we know about this topic?

Atopic dermatitis is a disease with a strong genetic component. Previous genome-wide association studies have identified several population-specific single genetic variations, but lack of functional characterization has led to a gap in understanding its heritability.

How does this study impact our current understanding and/or clinical management of this topic?

This paper demonstrates the involvement of a deep intronic SNP at 9q21.11 (rs7872806) influencing *tight junction protein 2* expression through methylation and contributing to atopic dermatitis presentation. The identified gene is a target for epigenetic therapy.

doi: 10.18176/jiaci.0978

Introduction

Atopic dermatitis (AD) is a chronic inflammatory itchy skin condition that affects more than 2% of the global population worldwide. In Singapore, this allergic disease affects about 13% of people [1] and decreases the quality of life of sufferers and caregivers [2]. It also poses a significant economic burden on Singapore's healthcare system, with its prevalence projected to increase [3].

AD is associated with a dysfunction in the skin epidermal barrier and results in the elevation of immunoglobulin E (IgE) levels in response to allergens. The pathogenesis of AD is multifactorial, influenced by both genetic and environmental factors. While the exact etiology and interplay of these factors remains not fully known, genetic predisposition to it has been shown through family and twin studies. In the European population, the AD concordance rate for monozygotic twins is 0.15 to 0.86, which is about 2 to 3 times higher as compared to the 0.05 to 0.41 for dizygotic twins [4]. Individuals with a family history of AD also present with a 2.5-fold increase in AD susceptibility [5]. Heritability for AD is estimated to be 71 to 84% [6].

As underscored by insights from Armario-Hita et al. (2023), understanding the disease's genetic foundations will empower personalized therapy, thereby maximizing treatment efficacy. Advancements in genome-wide genotyping arrays have driven the widespread adoption of genome-wide association studies (GWAS) in the quest to identify disease-associated genetic variants. Studies performed in populations of various ethnicities have identified more than 23 AD susceptibility loci to date (GWAS catalog: https://www.ebi.ac.uk/gwas/). Seven risk loci hypothesized to be involved in epidermal barrier function and immune dysregulation were identified in the GWAS of European descent individuals [8,9], two in Han Chinese [10], and two in the German population [11]. Twelve

doi: 10.18176/jiaci.0978

susceptibility loci were also identified in the Japanese population [12,13]. While some AD susceptibility loci were consistent across populations, a large proportion remained population specific. This can be attributed to the genetic heterogeneity of the disease between populations of varying ancestry groups or the presence of background-specific rare single nucleotide polymorphisms (SNPs) (minor allele frequencies less than 5%) with very strong disease effect. Hence, current GWAS identified loci are inadequate in comprehensively describing AD pathogenic heritability.

Here we performed a GWAS for AD in the Singapore Chinese population where we newly identified a SNP at 9q21.11 (rs7872806) strongly associated with the presence of AD. Through the integration of genetic and epigenetic association studies, methylation and expression quantitative trait loci (mQTL, eQTL), *in vitro* allele-specific luciferase expression effects and phenotypic evaluations, we attempt to functionally characterize this identified deep intronic SNP and illustrate its involvement in the pathogenesis of AD.

Materials and methods

Subject recruitment and Population Ascertainment

Study subjects of self-reported Chinese ethnicity residing in Singapore were recruited from the National University of Singapore (NUS) through various ongoing cross-sectional volunteer recruitment drives since August 2005. Approvals were obtained from the Institutional Review Board of NUS (IRB, Reference - NUS-07-023, NUS-09-256, NUS-10-445, NUS-13-075, NUS-14-150, and NUS-18-036) in compliance with the Helsinki Declaration together with written informed consent. Inclusion criteria were aged 18 and above, Chinese ethnicity, no consumption of antihistamine at least 3 days prior to the study and no previous participation by an immediate family member.

Subjects were aged around 21.46 years old (SD = 4.62) and classified into AD case or control as per validated guidelines from the UK Working Party Criteria, Hanifin and Rajka diagnostic criteria and ISAAC validated questionnaires [14–16]. The clinical symptoms of AD were defined to be a recurrent flexural itch lasting six months or longer. A house dust mite (HDM) skin prick test (SPT) was performed alongside to define atopy status, which showed high sensitivity and specificity in previous studies when compared to the clinical standard employing presence of allergic co-morbidities [17]. HDM skin prick test was used as the sole method to assess atopy based on the findings by Andiappan et al., (2014). In their study of the Singapore population, over 80% of IgE-sensitized participants tested positive for HDM-specific IgE. Among those who did not exhibit HDM-specific IgE, only 5% were atopic to other allergens and even so, their IgE titers were low [18]. AD cases were thus defined as subjects with AD symptoms and a positive atopic status. Atopic and non-atopic controls have a positive and negative atopic status respectively with no AD symptoms.

Genotyping and Genome Wide Association Studies

Genomic DNA was extracted from volunteers' buccal cells using Axygen® AxyPrep™ Multisource Genomic Miniprep DNA kit and subsequently genotyped using four GWAS arrays (Infinium OmniZhongHua-8 v1.3 BeadChip platform, Illumina HumanHap 550 k BeadChip, version 3, InfiniumOmni2-5Exome and Infinium Global Screening Array). In accordance with the recommendation put forth by Roshyara et al. (2014), advising imputation before quality control, we proceeded with haplotype phasing and imputation using IMPUTE v2.0 program with information from the 1000 Genomes Project phase III Beijing Han Chinese (CHB) database [20] which was previously demonstrated to possess high minor allele frequency (MAF) concordance with the Singapore Chinese [21]. Results from all 4 arrays were then combined.

doi: 10.18176/jiaci.0978

The samples were then subjected to quality control; removing individuals with a call rate < 90%, inconsistent gender information, duplicated and related samples. SNPs with call rates < 90%, indels, duplicated and monomorphic SNPs, and SNPs that did not pass the Hardy-Weinberg Equilibrium filter of $p > 0.001$ were also removed, yielding 5,215,687 merged SNPs. The detailed data preprocessing pipeline can be found in Supplemental figure S4.

Non-Chinese individuals were excluded following scrutiny of investigator-administered questionnaire responses, resulting in a cohort comprising 5,055 Singaporean individuals of Chinese descent. Given the clinical nature of AD, atopic controls were excluded from further GWAS analysis in order to minimise confounding from other allergic conditions and to more accurately represent individuals without inherent AD predisposition. Such exclusion also gives us a lower genomic inflation est. lambda of 1.04787. In total, we had 1679 unique participants (956 AD cases and 723 non-atopic controls) that were further selected for SNP-disease association analysis (Table S1).

Using PLINK program v1.09, statistical analysis was performed with logistic regression and adjusted for age, gender and 10 principal components (PCs). Manhattan and quantile-quantile plots were drawn using R program v3.6.2 qqman package. PLINK program v1.09 was used to identify independent AD association signals and calculate the genomic inflation factors (λ_{GC}), odds ratio (OR) and minor allele frequencies (MAF). SNPs with p-values $< 5 \times 10^{-8}$ (genome-wide significance threshold for this study) and MAF >5% were further examined. Haploview® version 4.2 was used to calculate the r^2 values, draw linkage disequilibrium (LD) patterns and determine haplotype frequencies for gene constructs [22]. The methodology used to obtain and select the principal components as well as power calculations can be found in the supplementary text.

doi: 10.18176/jiaci.0978

Transcriptome and Methylation Analysis in PBMCs

Whole blood was collected from 575 Singapore and Malaysia Chinese volunteers from the National University of Singapore (NUS), Sunway University or Universiti Tunku Abdul Rahman (UTAR) with approval from the Scientific and Ethical Review Committee (SERC) of UTAR (Ref-code: U/SERC/03/2016), Sunway University Research Ethics Committee (Ref - SUREC 2019/029) and Institutional Review Board at the National University of Singapore, Singapore (IRB numbers: NUS 07-023, NUS 09-256 and NUS 10-445). Total RNA of peripheral blood mononuclear cells (PBMCs) was extracted following isolation by Ficoll-Hypaque density gradient centrifugation using E.Z.N.A.® Total RNA kit from Omega Bio-tek Inc. (Norcross, GA), according to the manufacturer's instructions. mRNA was enriched using oligo(dT) beads and NEBNext Ultra RNA Library Prep Kit prior to sequencing using Illumina NovaSeq 6000 system. Raw sequences were mapped against the human genome assembly (GRCh37/hg19 version) using TopHat version 2.1.1. Fragments per kb of transcript per million mapped reads (FPKM) were calculated using Cufflinks version 2.2.1 where raw counts were normalised by sequencing depth and gene length. FPKM values were then quantile normalized. Transcript data (RNAseq) obtained was used for downstream expression quantitative trait loci (eQTL) analysis. Single-cell RNA sequencing (scRNA-seq) dataset "Fig2e-5PBMCs_scRNAseq_matrix.txt" obtained from CIBERSORTx© (Stanford University 2023) was used to create the signature matrix and obtain imputed cell fractions [23,24]. The imputed cell fractions were subsequently applied as covariates in a linear regression analysis to evaluate the impact of cell type heterogeneity in PBMCs on our RNAseq data. Association analysis was conducted using two-sample independent t-test with unequal variances. Transcript expression was also assessed using publicly available databases, namely eQTLGen Phase II (available at: https://www.eqtlgen.org/cis-eqtls.html) and GTExPortal version 8 (available at: https://gtexportal.org/).

doi: 10.18176/jiaci.0978

Genomic DNA was extracted from 233 out of the 575 Singapore and Malaysia Chinese PBMC samples and subjected to bisulphite conversion. Infinium MethylationEPIC BeadChip Kit (Illumina, San Diego, CA) was then used to assay the whole methylome. Annotation of the array data was conducted using the manifest file supplied by Illumina, titled "Infinium MethylationEPIC v1.0 B4 Manifest File," encompassing information for 865,918 probes (available at: https://sapac.support.illumina.com/downloads/infinium-methylationepic-v1-0-product-files.html). The raw data underwent preprocessing and quantile normalization, followed by the calculation of methylation beta-values, utilizing the minfi package within R version 3.6.1. To mitigate potential batch-related variations, all samples were concurrently analysed. Methylation levels at each CpG site 10kb up and downstream flanking the gene were evaluated for association with genotype via ANOVA with multiple testing corrected by Bonferroni adjustment or Pearson's correlation test. ReFACTor v1.0 was used to generate ReFACTor components, which were subsequently applied to assess the impact of cell type heterogeneity in PBMCs on the results of the methylation analysis through linear regression [25].

Cells and Treatments

Human embryonic kidney 293T cells (HEK293T) were purchased from American Type Culture Collection (ATCC, Manassas, VA) and grown in high-glucose Dulbecco's modified Eagle medium (DMEM, HyClone^{TM}; Cytiva) supplemented with 10% heat-inactivated fetal bovine serum (Gibco). Cells were cultured in a humidified incubator at 37°C with 5% CO_2, and passaged using trypsin (HyClone^{TM}; Cytiva) when they reach 90% confluency, up to 20 passages.

doi: 10.18176/jiaci.0978

The *TJP2* gene promoter region from 69,119,282 - 69,121,282 bp (Chr9, GRCh38/hg38) of the top 2 most frequently occurring haplotypes in the population were cloned into promoter-less pGL4.10 plasmids with firefly luciferase reporter gene (Promega, Singapore). Cells were seeded at a density of 2.5 x 10^5 cells in 24-well plates and cultured for 24 hours followed by replacement with fresh cell culture medium before transfection. When evaluating for methylation inhibition effect, media was replaced with fresh medium containing 0 µM, 0.1 µM, 0.5 µM or 1.0 µM of 5-aza-2'-deoxycytidine (DAC) instead, prior to transfection with reporter constructs.

HEK293T cells were transiently transfected at 70-80% confluency using Lipofectamine 2000 reagent (Invitrogen, Singapore) with 2 µg of reporter constructs according to the manufacturer's protocol. Four nanograms of control plasmid pGL4.74 with renilla luciferase reporter gene was also co-transfected to normalize for transfection efficiency. Cells were harvested 24 hours after transfection. Luciferase activity was measured using Promega's GloMax® Discover Microplate Reader and Dual Luciferase Reporter Assay Kit (Promega, Singapore) according to manufacturer's instructions. Firefly luciferase activities were normalised to that of Renilla to obtain the relative luciferase units (RLU). Transfection with empty vector pGL4.10 [luc] was also included as a control. All experiments were performed in triplicates.

Physiological Skin Parameters

Water barrier function (trans-epidermal water loss, TEWL) data were obtained using the Tewameter® TM300 (Courage+Khazaka electronic GmbH, Germany) respectively for 72 individuals from the Singapore and Malaysia Chinese cohort at the antecubital fossa of both arms following 20 minutes of acclimatization according to manufacturer's instructions. The average of triplicate measurements was used for downstream analysis.

doi: 10.18176/jiaci.0978

Results

Independent genomic signal at 9q21.11 (rs7872806) associated with atopic dermatitis

GWAS analysis of 956 AD cases and 723 non-atopic controls common SNP variants (i.e., those with MAF≥ 0.05) gave a λ_{GC} of 1.048. Twenty-six independent association signals at the stringent genome-wide significance threshold (GWAS $p < 5 \times 10^{-8}$) were identified (Figure 1A and B, Table 1), from which rs7872806 minor allele "A" was found to have the strongest association with a reduced risk of AD (adjusted p-value = 2.40×10^{-18}, OR = 0.38). Its minor allele frequency of 0.20 in our sampled population is similar to that reported in East Asians from the 1000 Genomes project. LD analysis identified 21 other SNPs with strong pairwise correlation (LD: $r^2 > 0.75$) to SNP rs7872806. These SNPs were located in either intronic or promoter regions of *TJP2* transcript variant 5, GenBank accession: NM_001170414.2 (Figure 2A and B, Table S2).

TJP2 SNP rs7872806 associated with higher TJP2 expression in PBMCs but not in skin epidermis

Given that the SNP rs7872806 lies in the deep intronic region of the *TJP2* gene (Figure 2A), we next examined if it had any effect on *TJP2* gene expression. On GTExPortal version 8, which contains *TJP2* expression data in various human tissues, we found SNP rs7872806 to significantly influence *TJP2* transcript expression only in whole blood (GTExPortal version 8 of 670 PBMCs, $p = 1.90 \times 10^{-7}$) but not in the human skin epidermis. This is despite expression being much higher in the skin (GTExPortal version 8, Figure S1). To verify this, we analysed *TJP2* transcript expression in PBMCs of 575 of the genotyped participants. Significantly higher *TJP2* transcripts (FPKM) was found with increasing number of the minor allele "A" (Figure 3A, T-test p <0.01). This differential expression pattern in PBMCs was likewise reported in

doi: 10.18176/jiaci.0978

eQTLGen Phase II database of 30,942 PBMCs (p = 1.05 x 10^{-119}). Similarly, no disease-*TJP2* transcript expression association was observed in skin biopsies from three independent studies [26–28](Figure S2). Covariates were not included in the analysis as their incorporation did not influence the significance of our results.

Based on further literature search, we found that the *TJP2* protein had alternative functions in intracellular signalling. *TJP2* was described to be a chief regulator of several processes controlling cell proliferative rates, size and number. This is through associating with nuclear protein Activator protein 1 (AP-1) that regulates *cyclin D1* (*CD1*) transcription [29,30]. We thus investigated the *AP-1* and *CD1* transcript levels in our PBMC samples and found an inverse relationship to *TJP2* expression concordant with our hypothesis (T-test $p < 0.001$, Figure S3).

TJP2 SNP rs7872806 associated with reduced methylation of CpG site cg13920460

Since SNP rs7872806 is deep intronic and lies >37 kb from the start of the *TJP2* gene, it is unlikely to influence splicing or transcription factor binding. Nevertheless, previous studies have demonstrated a link between genotypic variation in non-coding regions and the regulation of DNA methylation and disease, majority of which being cis-effects [31–33]. We thus investigated whether allele-specific methylation could account for the observed differential expression. This was performed through analysing the DNA methylation levels of CpG sites in the 10kb region flanking the *TJP2* gene, for association with rs7872806 genotype.

After Bonferroni correction for multiple testing ($p < 0.05/67$ CpG sites), of the 6 CpG sites with significant associations to rs7872806 genotype, only CpG site cg13920460 located 497bp upstream of the *TJP2* gene (Figure 3B, blue arrow, p = 2.12 x 10^{-11}) showed significant negative correlation with *TJP2* expression in PBMCs (Pearson's p = 1.09 x 10^{-6}, Pearson's R = -0.313, Figure 3C). The minor allele "A" of rs7872806 was significantly associated with reduced

doi: 10.18176/jiaci.0978

methylation at CpG site cg13920460 (T-test p <0.001) and higher *TJP2* expression as compared to the major allele (Figure 3C, 3D). The findings demonstrate robustness, even when a wider methylation window is applied, and when covariate and cell heterogeneity controls are included (data not shown).

CpG site cg13920460 in vitro methylation effect on TJP2 expression

We further sought to validate the effect of methylation on *TJP2* promoter activity *in vitro* by performing a luciferase assay. Three SNPs were identified to differ between the two most frequently occurring *TJP2* promoter haplotypes in the Singapore Chinese population - Haplotype 1 tagged to rs7872806 "G" major risk allele (58.5% frequency) and Haplotype 2 tagged to rs7872806 "A" minor protective allele (24.8% frequency) (Figure 4). Plasmid constructs were generated for both haplotypes (region Chr 9: 69,119,282 - 69,121,282 bp, GRCh38/hg38) and cloned into promoter-less pGL4.10 luciferase reporter gene plasmids. In 24 hours post-transfected HEK293T cells, basal RLU expression levels were very low but exposure to DAC treatment inhibiting *in vitro* methylation led to a significant increase in the activity of Haplotype 1 (G allele) promoter with dose response. Hence confirming the involvement of DNA methylation in regulating *TJP2* expression (T-test $p < 0.05$, Figure 5A). To determine whether the identified CpG site cg13920460 is consequential in regulating *TJP2* promoter activity, CpG site cg13920460 was deliberately eliminated through replacing the Cytosine base with a Thymine, cloned into another pGL4.10 plasmid and similarly transfected into HEK293T cells. The elimination of the single CpG site saw a significant increase in RLU expression levels of Haplotype 1 (G allele) by about 25%, to levels resembling that of Haplotype 2 (A allele). This signifies that CpG site cg13920460 plays a considerable role in regulating *TJP2* expression in mammalian cells through methylation (T-test $p < 0.01$, Figure 5B and 5C).

doi: 10.18176/jiaci.0978

SNP rs7872806 associated with reduced Trans-epidermal water loss in skin

Given *TJP2*'s predicted function in influencing cell proliferation rates, we hypothesized that it contributes to the epidermal hyperplasia phenotype commonly observed in AD patients. Studies performed by Pierce et al, 1998 and Robles et al, 1996 have also showed that overexpression of *CD1* led to the presentation of epidermal hyperproliferation and hyperplasia in the skin of transgenic mice. Our hypothesized pathway of how SNP rs7872806 confers AD risk and begets the pathogenic phenotype is shown in Figure 7. Since studies have previously demonstrated the association between epidermal hyperplasia and significantly higher trans-epidermal water loss (TEWL) levels [36–38], we employed the use of TEWL measurements as a non-invasive approach to envisage the presence of skin barrier dysfunction.

Physiological skin TEWL levels of 72 participants were measured at the antecubital fossa, a region commonly implicated in the disease. In concordance with our data, individuals carrying the minor allele "A" at SNP rs7872806 which was associated with higher *TJP2* expression, had significantly lower TEWL levels as compared to those with the major "G" allele (T-test $p < 0.001$, Figure 6).

Discussion

Previous genome-wide association studies have identified several AD-associated variants in populations of different ancestry. However, given the heterogeneity of the disease, several risk variants have yet to be identified and functionally characterized. The functional roles of deep intronic variants are also difficult to elucidate. In this study we demonstrate the epigenetic control of an identified deep intronic 9q21.11 polymorphism (SNP s7872806) in influencing *TJP2* expression and AD disease.

doi: 10.18176/jiaci.0978

The *TJP2* SNP rs7872806 is strongly associated with AD in our Singapore Chinese cohort, where its minor allele "A" conferred a protective effect against the disease by 2.64-fold. Given that the *TJP2* protein is expressed at high levels in the skin, its typically hypothesized function related to AD would have been structural. The tight junction proteins, also known as zonula occludens (ZO), are a member of the family of membrane-associated guanylate kinase homologs that serve as a scaffold for assembly of other tight junction proteins, bridging the connection between claudins and the actin cytoskeleton in the cytoplasm [39].

Investigations into *TJP2*'s tissue-specific expression in GTExPortal version 8 however, we found an interesting association with rs7872806 genotype only in whole blood but not in the skin (Figure S1). Similarly, in database eQTLGen Phase II of 30,942 PBMCs, the same differential expression pattern was reported in whole blood, but not in the skin. This was also replicated in the PBMCs extracted from our own participants, where subjects carrying the minor allele "A" had higher *TJP2* transcript expression (Figure 3A). We thus performed further literature search into TJP2's other alternative functions and found it to also be involved in intracellular signalling. The association between signalling molecules and AD is not a novel one, as demonstrated by studies such as Gamez et al. (2022) which established a connection between elevated levels of epithelial cell-derived cytokines to the early onset of AD. In a review by González-Mariscal et al. (2019), *TJP2* was described to be a chief regulator of several processes controlling cell proliferative rates, size and number through associating with nuclear proteins. In a study conducted by Betanzos et al. (2004), overexpression of *TJP2* in Madin-Darby canine kidney (MDCK) cells led to a decrease in *AP-1* controlled promoter activities in a dose-dependent manner. These sites, when bound by *AP-1* transcription factors, can result in upregulation of downstream proteins such as CD1. AP-1 binds two regulatory regions in the promoter of *CD1* [43] and induces its transcription, thereby promoting cell cycle G1 progression and cell proliferation [44,45]. Based on the understanding that patients with AD

doi: 10.18176/jiaci.0978

skin present with a significantly higher index of proliferative cells as compared to healthy controls [46], we hypothesize that *TJP2*'s effect on AD phenotype might be through AP-1 and CD-1. This was supported by studies performed by Girolomoni and Pastore (2001) where *AP-1* activation was found to be upregulated in AD keratinocytes and by Kim et al. (2018) which found higher *CD1* expression in chronic psoriatic lesions as compared to in control skin, a disease presenting with a similar phenotype as AD. Another study by Mao et al. (2007) also identified that the Cyproheptadine drug commonly used to treat AD had a inhibitory effect on *CD1* expression. *In vitro* studies investigating the knockdown of *TJP2* in MDCK cells have also found an increase in *CD1* expression [29]. From the PBMC samples of our participants, we also observed a significant inversely proportional relationship between *TJP2* and *AP-1* or *CD1* expression (Figure S3).

Putting these findings together we hypothesize that low levels of *TJP2* protein act to increase the activity of *AP-1* and subsequently the transcriptional activity of *AP-1* regulated promoters, one of which is *CD1*. Thereby the increase in *CD1* expression results in abnormally elevated cell proliferation rates and numbers giving rise to a thickened skin epidermis, a condition known as epidermal hyperplasia, which is a characteristic phenotype seen in AD patients (Figure 7). This relationship between elevated CD1 levels and epidermal hyperproliferation was also demonstrated in a study by Robles et al. (1996) where *CD1* was overexpressed in the skin of transgenic mice. Given that several studies have demonstrated an association between the presence of epidermal hyperplasia and high TEWL levels in the skin [36–38], we measured TEWL levels on the skin of participants as a non-invasive approach to determining the existence of skin barrier dysfunction and epidermal hyperplasia. From the data collected we were able to show that individuals carrying the major "G" allele of SNP rs7872806 had significantly higher antecubital fossa TEWL levels, with none of the individuals carrying the "A" allele presenting with TEWL levels above 23.19 $g \cdot m^{-2} \cdot h^{-1}$ (Figure 6), a cut off previously

doi: 10.18176/jiaci.0978

determined by Montero-Vilchez et al. (2021) for moderate and severe AD with 73.5% sensitivity.

Understanding that deep intronic SNPs such as rs7872806 are unlikely to influence splicing [51], we investigated whether it had any cis-regulated allele-specific methylation effect instead, which has been demonstrated in other studies investigating complex diseases [31–33]. This led to the discovery of associations between SNP rs7872806 genotype and differential methylation levels at the CpG site cg13920460 located in the promoter region of *TJP2* gene (Figure 3B), and which was significantly inversely correlated with *TJP2* gene expression in PBMCs (Figure 3C). The methylation effect was further validated through *in vitro* methylation inhibition assays using DAC in HEK293T cells with dose response (Figure 5A). Elimination of the single CpG site cg13920460 was also sufficient to significantly raise the RLU (Figure 5B and 5C), indicating a considerable role of this site in regulating *TJP2* expression. The complete hypothesized pathogenic pathway is thus illustrated in Figure 7.

To gain further insights of our genetic variant, we explored the FUMA GWAS platform (https://fuma.ctglab.nl/snp2gene) and noted a publication by Ter Hark et al., (2020) on SNP rs7872806's LD variant rs7024062. The authors reported a significant association between rs7024062's major allele and antipsychotic drug-induced weight gain (p value = 2.2×10^{-6}). In other case studies such as the one by Bujor et al., (2017), antipsychotic-treated patients that experience weight gain also developed psoriatic lesions. This suggests that the major allele of rs7024062 may also be associated with psoriatic lesion formation. Considering that epidermal hyperplasia is a hallmark of psoriatic skin lesions, the literature findings further strengthen our hypothesis that SNP rs7872806's major allele contributes to the AD pathological pathway through epidermal hyperplasia. Despite conducting an extensive literature search, this was the only scientific literature pertaining to the leading variant or its LD variants. This scarcity of research highlights the limited exploration of this specific genetic variant in existing literature.

doi: 10.18176/jiaci.0978

While the other SNPs in LD with rs7872806 exhibited identical patterns of *TJP2* transcript expression and allele-specific methylation as rs7872806 (data not shown) due to their high LD (LD: $r^2 > 0.75$), we hypothesize that rs7872806 is the SNP with the highest likelihood of being causal. This is because the GWAS association p-value for rs7872806 was over 10^{14}-fold more significant in relation to the disease compared to its LD variants, strongly indicating its potential role as the causal site or the site with the most substantial effect size. However, we acknowledge that definitive confirmation of rs7872806 as the causal site would require further targeted *in vitro* perturbations since our *in vitro* studies currently evaluate methylation effects as a haplotype.

Our study, nonetheless, does present with some limitations, one of which is the absence of replication. Following the approach advocated by Skol et al., (2006), which underscores the advantages of joint data analysis over replication-based methods due to its higher statistical power, we opted to pool all available samples for our GWAS. To address the need for replication, we conducted an in-silico search of various public GWAS databases (UK Biobank, Finngen, Biobank Japan Pheweb, MGL PheWeb). Unfortunately, we did not find any prior reports supporting the significance of rs7872806 in these datasets. For instance, in the Biobank Japan Pheweb GWAS, which included 4,296 atopic dermatitis cases and 163,807 controls, the SNP rs7872806 yielded a *p*-value of 6.3×10^{-1} with a minor allele effect size of -0.015 [55]. Similarly, the MGL PheWeb GWAS, encompassing 5,018 atopic/contact dermatitis cases and 35,760 controls, produced a *p*-value of 8.9×10^{-1} and a minor allele effect size of -0.0041 for SNP rs7872806, none of which reached GWAS significance. This outcome is not entirely unexpected, given the genetic heterogeneity across populations of different ancestries. Replication in diverse populations is therefore challenging. Moreover, our proposed pathological pathway involves epigenetic mechanisms, which are strongly influenced by

doi: 10.18176/jiaci.0978

various environmental factors that vary among populations. Future replication studies within the same population could offer further validation of our results.

Another aspect of our study that may raise concerns is the relatively smaller size of our control group that consists of 723 individuals, compared to the 956 cases. The smaller control group size was a result of our strict inclusion criteria for control participants. However, even if we were to expand the control group to include atopic individuals, increasing the control sample size to 2,961, the result was an even more significant GWAS association p-value of 5.40×10^{-39} for SNP rs7872806. This finding underscores that the smaller size of the original control group did not compromise the validity of our study. Nonetheless, we acknowledge the potential benefit of further increasing both the case and control sample sizes in future investigations to enhance the robustness of our findings.

In our *in vitro* expression studies, we made the unconventional choice of employing HEK293T cells due to the specific context and requirements of our study. The traditionally favoured cell model for AD are keratinocytes, them being the primary cell type involved in AD-related skin barrier dysfunction and inflammation. However, since our investigation centres around the signalling function of the *TJP2* gene rather than its structural role, the observed effects need not be replicated in keratinocytes. Additionally, our early investigations of eQTL in keratinocytes from GTExPortal version 8 did not reveal significant differential *TJP2* expression associated with rs7872806 genetic variants in the skin. The use of PBMCs as a cell model was also ruled out due to their heterogeneity, comprising a mix of various cell types. During our literature search, we came across a study conducted by Betanzos et al., (2004) that utilized Madin-Darby canine kidney (MDCK) cells as a model to investigate *TJP2* overexpression. This provided a precedent for exploring *TJP2* expression effects in epithelial-like kidney cells like HEK293T, which are of human origin. Moreover, HEK293T cells offer distinct advantages, including high transfection efficiency and proficient protein expression

doi: 10.18176/jiaci.0978

capabilities, attributes that make them ideal for gene expression experiments. In future research endeavours, further evaluation of *TJP2* expression in other cell models may offer a more comprehensive understanding of AD pathogenesis.

In addition to the results presented, we also conducted other analyses, but these failed to reveal significant associations. For instance, we explored potential correlations between disease phenotype and *TJP2* expression or cg13920460 methylation, but our findings did not reach statistical significance. This may be due to the limited case sample size used for the RNAseq and methylation studies, coupled with the heterogenous and complex nature of AD, which can obscure direct associations. Nevertheless, the consistent findings from other evidence sources still provides confidence and reaffirm our overall conclusions. Expanding the cohort size in future work could enhance our ability to detect significance in these associations.

We also ventured to investigate whether there was a correlation between the frequency of rs7872806's minor allele 'A' (retrieved from NCBI) and reported AD prevalence in various human populations [56,57], as documented in Supplementary Table S3. Our analysis however, yielded no correlation. Nevertheless, this lack of significant correlation aligns with our hypothesis that the pathological pathway of SNP rs7872806 is mediated through methylation, a process highly influenced by environmental factors that varies across populations.

In summary, this study has identified and characterised a possible mechanism for the functional role of 9q21.11 polymorphism in contributing to AD phenotype. The SNP rs7872806 increases methylation at CpG site cg13920460 thereby influencing *TJP2* expression which regulates downstream *AP-1* and *CD1* levels resulting in increased cell proliferation and trans-epidermal water loss suggesting epidermal hyperplasia development. Our findings also provide insights into the epigenetic control of non-coding genetic polymorphisms and its ability to influence complex disease phenotypes. Such epigenetic mechanisms can thus be further investigated as potential targets for therapeutic intervention.

doi: 10.18176/jiaci.0978

Acknowledgements

We extend our sincerest gratitude to all participants for their contributions to this study. Figure 7 created with BioRender.com

Funding

F.T.C. has received research support from the National University of Singapore, Singapore Ministry of Education Academic Research Fund, Singapore Immunology Network (SIgN), National Medical Research Council (NMRC) (Singapore), Biomedical Research Council (BMRC) (Singapore), National Research Foundation (NRF) (Singapore), Singapore Food Agency (SFA), and the Agency for Science Technology and Research (A*STAR) (Singapore); Grant Numbers are N-154-000-038-001, R-154-000-191-112, R-154-000-404-112, R-154-000-553-112, R-154-000-565-112, R-154-000-630-112, R-154-000-A08-592, R-154-000-A27-597, R-154-000-A91-592, R-154-000-A95-592, R-154-000-B99-114, BMRC/01/1/21/18/077, BMRC/04/1/21/19/315, BMRC/APG2013/108, SIgN-06-006, SIgN-08-020, NMRC/1150/2008, OFIRG20nov-0033, NRF-MP-2020-0004, SFS_RND_SUFP_001_04, W22W3D0006, H17/01/a0/008 and APG2013/108. F.T.C. has received consulting fees from Sime Darby Technology Centre, First Resources Ltd, Genting Plantation, Olam International and Syngenta Crop Protection, outside the submitted work. Y.Y.S. has received research support from the NUS Resilience & Growth Postdoctoral Fellowships with grant number R-141-000-036-281. Y.Y.E.L. has received funding from the National University of Singapore Department of Biological Sciences President's Graduate Fellowship. All funding agencies had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosure of competing interests

The authors declare no competing interests.

Abbreviations

Activator protein 1, (AP-1); Atopic dermatitis, (AD); American Type Culture Collection, (ATCC); Beijing Han Chinese, (CHB); Cyclin D1, (CD1); 5-aza-2'-deoxycytidine, (DAC); Dulbecco's Modified Eagle Medium, (DMEM); Expression quantitative trait loci, (eQTL); Fragments Per Kilobase of transcript per Million mapped reads, (FPKM); Genome-wide association study, (GWAS); Human embryonic kidney 293T cells, (HEK293T); House dust mite, (HDM); Immunoglobulin E, (IgE); Institutional review board, (IRB); International Study of Asthma and Allergies in Childhood, (ISAAC); Linkage disequilibrium, (LD); Madin-Darby canine kidney, (MDCK); Minor allele frequencies, (MAF); Methylation quantitative trait loci, (mQTL); National University of Singapore, (NUS); Odds ratio, (OR); Principal components, (PCs); Peripheral blood mononuclear cells, (PBMCs); Relative luciferase units, (RLU); S-Adenosyl methionine, (SAM); Scientific and Ethical Review Committee, (SERC); Single nucleotide polymorphisms, (SNPs); Skin-prick test, (SPT); Trans-epidermal water loss, (TEWL); Tight junction protein 2, (TJP2); Universiti Tunku Abdul Rahman, (UTAR); Zonula occludens, (ZO)

References

1. Cheok S, Yee F, Song Ma JY, Leow R, Ho MSL, Yew YW, et al. Prevalence and descriptive epidemiology of atopic dermatitis and its impact on quality of life in Singapore [Internet]. Vol. 178, British Journal of Dermatology. Blackwell Publishing Ltd; 2018 [cited 2020 Feb 18]. p. 276–7. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/bjd.15587
2. Silverberg JI, Gelfand JM, Margolis DJ, Boguniewicz M, Fonacier L, Grayson MH, et al. Patient burden and quality of life in atopic dermatitis in US adults: A population-based cross-sectional study. Ann Allergy, Asthma Immunol. 2018 Sep 1;121(3):340–7.
3. Lee BW, Detzel PR. Treatment of childhood atopic dermatitis and economic burden of illness in Asia Pacific Countries. Ann Nutr Metab [Internet]. 2015 [cited 2020 Mar 20];66(1):18–24. Available from: www.karger.com/anm
4. Elmose C, Thomsen SF. Twin Studies of Atopic Dermatitis: Interpretations and Applications in the Filaggrin Era. J Allergy [Internet]. 2015 [cited 2021 Apr 27];2015:1–7. Available from: http://dx.doi.org/10.1155/2015/902359
5. Matsuoka S, Nakagawa R, Nakayama H, Yamashita K, Kuroda Y. Prevalence of specific allergic diseases in school children as related to parental atopy. Pediatr Int [Internet]. 1999 Feb 1 [cited 2020 Mar 31];41(1):46–51. Available from: http://doi.wiley.com/10.1046/j.1442-200x.1999.01011.x
6. Ober C, Yao TC. The genetics of asthma and allergic disease: A 21st century perspective. Immunol Rev. 2011 Jul;242(1):10–30.
7. Armario-Hita JC, Galán-Gutiérrez M, Dodero-Anillo JM, Carrascosa JM, Ruiz-Villaverde R. Updated Review on Treatment of Atopic Dermatitis. Vol. 33, Journal of Investigational Allergology and Clinical Immunology. 2023. p. 158–67.
8. Paternoster L, Standl M, Chen CM, Ramasamy A, Bøpnnelykke K, Duijts L, et al.

Meta-analysis of genome-wide association studies identifies three new risk loci for atopic dermatitis. Nat Genet. 2012;44(2):187–92.

9. Ellinghaus D, Baurecht H, Esparza-Gordillo J, Rodríguez E, Matanovic A, Marenholz I, et al. High-density genotyping study identifies four new susceptibility loci for atopic dermatitis. Nat Genet. 2013 Jul 2;45(7):808–12.

10. Sun LD, Xiao FL, Li Y, Zhou WM, Tang HY, Tang XF, et al. Genome-wide association study identifies two new susceptibility loci for atopic dermatitis in the Chinese Han population. Nat Genet. 2011;43(7):690–4.

11. Schaarschmidt H, Ellinghaus D, Rodrıguez E, Anja K, Baurecht H, Lipinski S, et al. A genome-wide association study reveals 2 new susceptibility loci for atopic dermatitis. J Allergy Clin Immunol [Internet]. 2015 [cited 2023 Sep 19];136(3):802–6. Available from: http://dx.doi.org/10.1016/j.jaci.2015.02.017

12. Hirota T, Takahashi A, Kubo M, Tsunoda T, Tomita K, Sakashita M, et al. Genome-wide association study identifies eight new susceptibility loci for atopic dermatitis in the Japanese population. Nat Genet. 2012 Nov;44(11):1222–6.

13. Tanaka N, Koido M, Suzuki A, Otomo N, Suetsugu H, Kochi Y, et al. Eight novel susceptibility loci and putative causal variants in atopic dermatitis. J Allergy Clin Immunol. 2021 Nov 1;148(5):1293–306.

14. Asher MI, Keil U, Anderson HR, Beasley R, Crane J, Martinez F, et al. International study of asthma and allergies in childhood (ISAAC): Rationale and methods. In: European Respiratory Journal. European Respiratory Society; 1995. p. 483–91.

15. Williams HC, JBURNEY PG, HAY RJ, ARCHER CB, SHIPLEY MJ, AHUNTER JJ, et al. The U.K. Working Party's Diagnostic Criteria for Atopic Dermatitis. I. Derivation of a minimum set of discriminators for atopic dermatitis. Br J Dermatol [Internet]. 1994 [cited 2023 Jul 7];131(3):383–96. Available from:

https://pubmed.ncbi.nlm.nih.gov/7918015/

16. Hanifin JM, Rajka G. Diagnostic Features of Atopic Dermatitis. Acta Derm Venereol [Internet]. 1980 Nov 11;60(92):44–7. Available from: https://medicaljournalssweden.se/actadv/article/view/10725

17. Lim JJ, Lim YYE, Ng JY, Malipeddi P, Ng YT, Teo WY, et al. An update on the prevalence, chronicity, and severity of atopic dermatitis and the associated epidemiological risk factors in the Singapore/Malaysia Chinese young adult population: A detailed description of the Singapore/Malaysia Cross-Sectional Genetics Epidemiology Study (SMCGES) cohort. World Allergy Organ J. 2022 Dec 1;15(12):100722.

18. Andiappan AK, Puan KJ, Lee B, Nardin A, Poidinger M, Connolly J, et al. Allergic airway diseases in a tropical urban environment are driven by dominant mono-specific sensitization against house dust mites. Allergy Eur J Allergy Clin Immunol [Internet]. 2014 [cited 2023 Sep 12];69(4):501–9. Available from: https://onlinelibrary.wiley.com/doi/10.1111/all.12364,

19. Roshyara NR, Kirsten H, Horn K, Ahnert P, Scholz M. Impact of pre-imputation SNP-filtering on genotype imputation results. BMC Genet [Internet]. 2014 Aug 12 [cited 2023 Sep 19];15(1):1–11. Available from: https://bmcgenomdata.biomedcentral.com/articles/10.1186/s12863-014-0088-5

20. Genomes Project C, Auton A, Abecasis GR, Altshuler DM, Durbin RM, Bentley DR, et al. A global reference for human genetic variation. Vol. 526, Nature. Nature Publishing Group; 2015. p. 68–74.

21. Andiappan AK, Anantharaman R, Nilkanth PP, Wang DY, Chew FT. Evaluating the transferability of Hapmap SNPs to a Singapore Chinese population. BMC Genet. 2010 May 7;11.

22. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics [Internet]. 2005 Jan 15 [cited 2023 Jul 27];21(2):263–5. Available from: https://dx.doi.org/10.1093/bioinformatics/bth457

23. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol [Internet]. 2019 May 6 [cited 2023 Oct 18];37(7):773–82. Available from: https://www-nature-com.libproxy1.nus.edu.sg/articles/s41587-019-0114-2

24. Steen CB, Liu CL, Alizadeh AA, Newman AM. Profiling cell type abundance and expression in bulk tissues with CIBERSORTx. In: Methods in Molecular Biology [Internet]. NIH Public Access; 2020 [cited 2023 Oct 17]. p. 135–57. Available from: /pmc/articles/PMC7695353/

25. Rahmani E, Zaitlen N, Baran Y, Eng C, Hu D, Galanter J, et al. Sparse PCA corrects for cell type heterogeneity in epigenome-wide association studies. Nat Methods [Internet]. 2016 May 1 [cited 2023 Oct 13];13(5):443–5. Available from: https://pubmed.ncbi.nlm.nih.gov/27018579/

26. Plager DA, Leontovich AA, Henke SA, Davis MDP, Mcevoy MT, Sciallis GF, et al. Early cutaneous gene transcription changes in adult atopic dermatitis and potential clinical implications. Exp Dermatol. 2007 Jan;16(1):28–36.

27. De Benedetto A, Rafaels NM, McGirt LY, Ivanov AI, Georas SN, Cheadle C, et al. Tight junction defects in patients with atopic dermatitis. J Allergy Clin Immunol [Internet]. 2011 [cited 2020 Apr 16];127(3):773-786.e7. Available from: www.jacionline.org

28. Suárez-Fariñas M, Tintle SJ, Shemer A, Chiricozzi A, Nograles K, Cardinale I, et al. Nonlesional atopic dermatitis skin is characterized by broad terminal differentiation

defects and variable immune abnormalities. J Allergy Clin Immunol [Internet]. 2011 Apr [cited 2022 Dec 7];127(4):954-64.e1-4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21388663

29. González-Mariscal L, Miranda J, Raya-Sandino A, Domínguez-Calderón A, Cuellar-Perez F. ZO-2, a tight junction protein involved in gene expression, proliferation, apoptosis, and cell size regulation. Ann N Y Acad Sci [Internet]. 2017 Jun 1 [cited 2022 Dec 9];1397(1):35–53. Available from: https://onlinelibrary-wiley-com.libproxy1.nus.edu.sg/doi/full/10.1111/nyas.13334

30. Huerta M, Muñoz R, Tapia R, Soto-Reyes E, Ramírez L, Recillas-Targa F, et al. Cyclin D1 Is Transcriptionally Down-Regulated by ZO-2 via an E Box and the Transcription Factor c-Myc. Mol Biol Cell [Internet]. 2007 Dec [cited 2022 Dec 9];18(12):4826. Available from: /pmc/articles/PMC2096592/

31. Meaburn EL, Schalkwyk LC, Mill J. Allele-specific methylation in the human genome: Implications for genetic studies of complex disease. Epigenetics [Internet]. 2010 Oct 27 [cited 2021 Apr 22];5(7):578–82. Available from: http://www.tandfonline.com/doi/abs/10.4161/epi.5.7.12960

32. Hutchinson JN, Raj T, Fagerness J, Stahl E, Viloria FT, Gimelbrant A, et al. Allele-Specific Methylation Occurs at Genetic Variants Associated with Complex Disease. Toland AE, editor. PLoS One [Internet]. 2014 Jun 9 [cited 2021 Apr 22];9(6):e98464. Available from: https://dx.plos.org/10.1371/journal.pone.0098464

33. Johnson MD, Mueller M, Adamowicz-Brice M, Collins MJ, Gellert P, Maratou K, et al. Genetic Analysis of the Cardiac Methylome at Single Nucleotide Resolution in a Model of Human Cardiovascular Disease. Mott R, editor. PLoS Genet [Internet]. 2014 Dec 4 [cited 2021 Apr 22];10(12):e1004813. Available from: https://dx.plos.org/10.1371/journal.pgen.1004813

34. Robles AI, Larcher F, Whalin RB, Murillas R, Richie E, Gimenez-Conti IB, et al. Expression of cyclin D1 in epithelial tissues of transgenic mice results in epidermal hyperproliferation and severe thymic hyperplasia. Proc Natl Acad Sci U S A. 1996 Jul 23;93(15):7634–8.

35. Pierce AM, Fisher SM, Conti CJ, Johnson DG. Deregulated expression of E2F1 induces hyperplasia and cooperates with ras in skin tumor development. Oncogene 1998 1610 [Internet]. 1998 Mar 19 [cited 2023 Feb 15];16(10):1267–76. Available from: https://www.nature.com/articles/1201666

36. Feingold KR, Mao-Qiang M, Proksch E, Menon GK, Brown BE, Elias PM. The lovastatin-treated rodent: A new model of barrier disruption and epidermal hyperplasia. J Invest Dermatol. 1991;96(2):201–9.

37. Imokawa G, Yada Y, Higuchi K, Okuda M, Ohashi Y, Kawamata A. Pseudo-acylceramide with linoleic acid produces selective recovery of diminished cutaneous barrier function in essential fatty acid-deficient rats and has an inhibitory effect on epidermal hyperplasia. J Clin Invest. 1994 Jul 1;94(1):89–96.

38. Fuziwara S, Suzuki A, Inoue K, Denda M. Dopamine D2-Like Receptor Agonists Accelerate Barrier Repair and Inhibit the Epidermal Hyperplasia Induced by Barrier Disruption. J Invest Dermatol. 2005 Oct 1;125(4):783–9.

39. Tsukita S, Furuse M. Pores in the Wall: Claudins Constitute Tight Junction Strands Containing Aqueous Pores. J Cell Biol [Internet]. 2000 Apr 4 [cited 2022 Jul 26];149(1):13. Available from: /pmc/articles/PMC2175101/

40. Gamez C, Metcalfe J, Prescott SL, Palmer DJ. Circulating Epithelial Cell Cytokines Are Associated With Early-Onset Atopic Dermatitis [Internet]. Vol. 32, Journal of Investigational Allergology and Clinical Immunology. 2022 [cited 2023 Sep 15]. p. 48–50. Available from: www.allergen.org

41. González-Mariscal L, Gallego-Gutiérrez H, González-González L, Hernández-Guzmán C. ZO-2 is a master regulator of gene expression, cell proliferation, cytoarchitecture, and cell size [Internet]. Vol. 20, International Journal of Molecular Sciences. 2019 [cited 2022 Dec 9]. Available from: www.mdpi.com/journal/ijms

42. Betanzos A, Huerta M, Lopez-Bayghen E, Azuara E, Amerena J, González-Mariscal L. The tight junction protein ZO-2 associates with Jun, Fos and C/EBP transcription factors in epithelial cells. Exp Cell Res [Internet]. 2004 Jan 1 [cited 2023 Mar 29];292(1):51–66. Available from: https://pubmed.ncbi.nlm.nih.gov/14720506/

43. Herber B, Truss M, Beato M, Müller R. Inducible regulatory elements in the human cyclin D1 promoter. Oncogene [Internet]. 1994 Apr 1 [cited 2023 Mar 29];9(4):1295–304. Available from: https://europepmc.org/article/med/8134134

44. Wisdom R, Johnson RS, Moore C. c-Jun regulates cell cycle progression and apoptosis by distinct mechanisms. EMBO J [Internet]. 1999 Jan 4 [cited 2023 Mar 29];18(1):188–97. Available from: https://onlinelibrary.wiley.com/doi/full/10.1093/emboj/18.1.188

45. Albanese C, Johnson J, Watanabe G, Eklund N, Vu D, Arnold A, et al. Transforming p21ras mutants and c-Ets-2 activate the cyclin D1 promoter through distinguishable regions. J Biol Chem [Internet]. 1995 [cited 2023 Mar 29];270(40):23589–97. Available from: https://pubmed.ncbi.nlm.nih.gov/7559524/

46. Sapuntsova SG, Mel'nikova NP, Deigin VI, Kozulin EA, Timoshin SS. Proliferative processes in the epidermis of patients with atopic dermatitis treated with thymodepressin. Bull Exp Biol Med [Internet]. 2002 May 1 [cited 2022 Dec 9];133(5):488–90. Available from: https://pubmed-ncbi-nlm-nih-gov.libproxy1.nus.edu.sg/12420069/

47. Girolomoni G, Pastore S. The role of keratinocytes in the pathogenesis of atopic

dermatitis. In: Journal of the American Academy of Dermatology. Mosby; 2001. p. S25–8.

48. Kim SA, Ryu YW, Kwon J Il, Choe MS, Jung JW, Cho JW. Differential expression of cyclin D1, Ki-67, pRb, and p53 in psoriatic skin lesions and normal skin. Mol Med Rep [Internet]. 2018 Jan 1 [cited 2022 Dec 9];17(1):735–42. Available from: http://www.spandidos-publications.com/10.3892/mmr.2017.8015/abstract

49. Mao X, Stewart AK, Hurren R, Gronda M, Lee K, Chow S, et al. A High Throughput Screen Identified Cyproheptadine That Decreases D-Type Cyclins, Arrests Cells in the G1 Phase, and Induces Apoptosis in Multiple Myeloma and Leukemia Cells. Blood. 2007 Nov 16;110(11):1507.

50. Montero-Vilchez T, Segura-Fernández-nogueras MV, Pérez-Rodríguez I, Soler-Gongora M, Martinez-Lopez A, Fernández-González A, et al. Skin Barrier Function in Psoriasis and Atopic Dermatitis: Transepidermal Water Loss and Temperature as Useful Tools to Assess Disease Severity. J Clin Med 2021, Vol 10, Page 359 [Internet]. 2021 Jan 19 [cited 2023 Mar 29];10(2):359. Available from: https://www.mdpi.com/2077-0383/10/2/359/htm

51. Woolfe A, Mullikin JC, Elnitski L. Genomic features defining exonic variants that modulate splicing. Genome Biol [Internet]. 2010 Feb 16 [cited 2021 Apr 22];11(2):R20. Available from: /pmc/articles/PMC2872880/

52. Ter Hark SE, Jamain S, Schijven D, Lin BD, Bakker MK, Boland-Auge A, et al. A new genetic locus for antipsychotic-induced weight gain: A genome-wide study of first-episode psychosis patients using amisulpride (from the OPTiMiSE cohort). J Psychopharmacol [Internet]. 2020 May 1 [cited 2023 Sep 15];34(5):524–31. Available from: https://doi.org/10.1177/0269881120907972

53. Bujor CE, Vang T, Nielsen J, Schjerning O. Antipsychotic-associated psoriatic rash - a

case report. BMC Psychiatry [Internet]. 2017 Jul 4 [cited 2023 Sep 28];17(1). Available from: /pmc/articles/PMC5496600/

54. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet [Internet]. 2006 [cited 2023 Oct 9];38(2):209–13. Available from: http://www.nature.com/naturegenetics

55. Sakaue S, Kanai M, Tanigawa Y, Karjalainen J, Kurki M, Koshiba S, et al. A cross-population atlas of genetic associations for 220 human phenotypes. Nat Genet 2021 5310 [Internet]. 2021 Sep 30 [cited 2023 Sep 18];53(10):1415–24. Available from: https://www-nature-com.libproxy1.nus.edu.sg/articles/s41588-021-00931-x

56. Hadi HA, Tarmizi AI, Khalid KA, Gajdács M, Aslam A, Jamshed S. The epidemiology and global burden of atopic dermatitis: A narrative review [Internet]. Vol. 11, Life. Multidisciplinary Digital Publishing Institute (MDPI); 2021 [cited 2023 Sep 13]. Available from: /pmc/articles/PMC8470589/

57. Bylund S, Von Kobyletzki LB, Svalstedt M, Svensson Å. Prevalence and incidence of atopic dermatitis: A systematic review. Acta Derm Venereol. 2020;100(100-year theme Atopic dermatitis):320–9.

FIGURE

Figure 1. GWAS for 956 Singapore Chinese atopic dermatitis cases and 723 non-atopic controls. **A**: Manhattan plot showing the -$\log_{10} P$ values of 5,215,687 SNPs in the cohort plotted against their respective positions on autosomes, after adjusting for age, gender and PC1-10. SNPs with MAF <5% were excluded. The red line indicates the genome-wide significance threshold for this study ($p = 5 \times 10^{-8}$). The blue line shows the GWAS suggestive threshold ($p = 1 \times 10^{-5}$). $\lambda_{GC} = 1.048$. Twenty six independent signals with GWAS $p < 5 \times 10^{-8}$ were obtained. Independent rs7872806 signal of interest labelled in red box. **B**: Quantile-quantile (Q-Q) plot of p-values observed plotted against expected. Red line indicates values predicted for a normal distribution. The probability plot illustrates that the GWAS data follows a normal distribution.

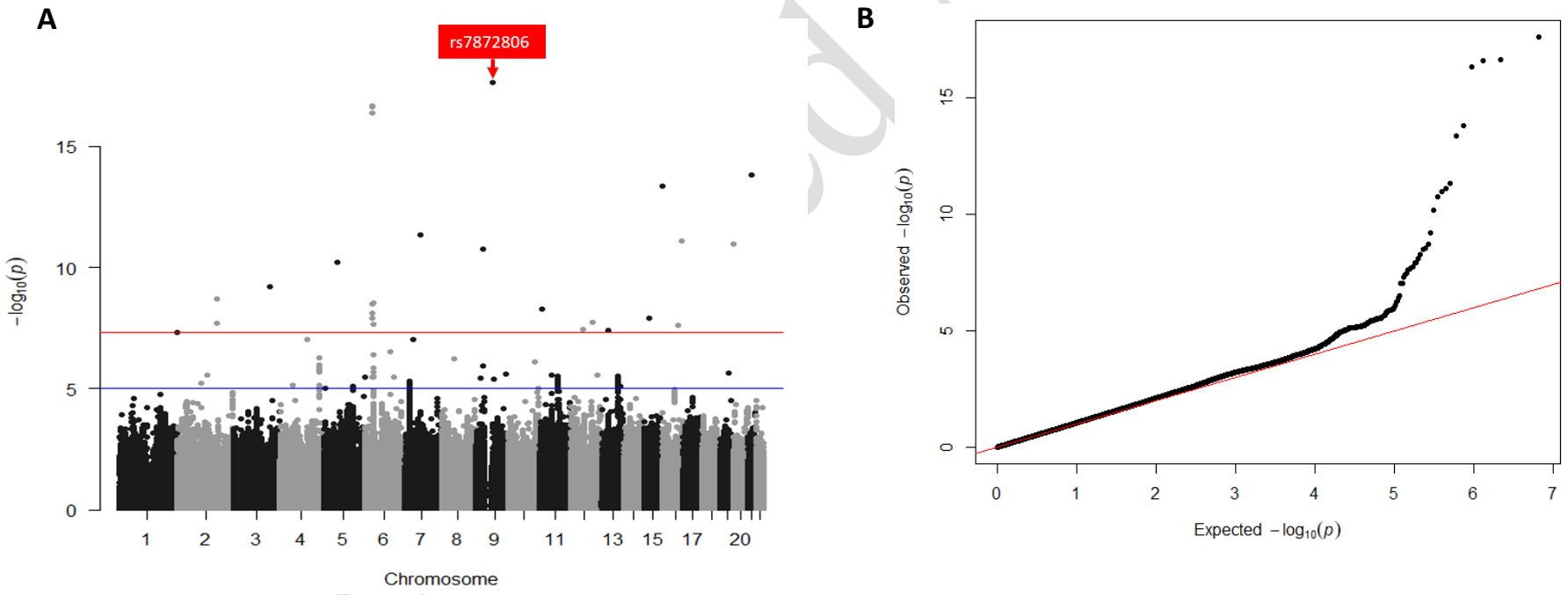

Table 1. List of the 26 AD-associated SNPs that pass GWAS significance threshold. The *p*-value is calculated using logistic regression.

CHR	SNP	P value	BP (GRCh37)	Minor/Major Allele	MAF (case/ctrl)	OR (95% CI)	Genes within 50kb of SNP
9	rs7872806	2.40E-18	71772940	A / G	0.09 / 0.20	0.38 (0.30 - 0.47)	*TJP2*
6	rs62391770	2.19E-17	29664983	T / C	0.07 / 0.17	0.36 (0.28 - 0.45)	*MOG; ZFP57; HLA-F; HLA-F-AS1*
6	rs148694836	2.41E-17	29665741	G / C	0.07 / 0.18	0.36 (0.29 - 0.46)	*MOG; ZFP57; HLA-F; HLA-F-AS1*
6	rs56269224	4.36E-17	29669998	G / C	0.07 / 0.18	0.37 (0.29 - 0.46)	*MOG; ZFP57; HLA-F; HLA-F-AS1; IFITM4P*
21	rs181286089	1.52E-14	27726488	A / G	0.03 / 0.10	0.27 (0.20 - 0.38)	*CYYR1-AS1*
15	rs1564638	4.39E-14	101895981	G / A	0.06 / 0.14	0.37 (0.29 - 0.48)	*PCSK6; LOC100507472; PCSK6-AS1*
7	rs34657769	4.68E-12	67366661	A / C	0.03 / 0.09	0.30 (0.21 - 0.42)	-
16	rs79224535	7.94E-12	84225184	T / C	0.02 / 0.08	0.26 (0.18 - 0.39)	*HSDL1; DNAAF1; TAF1C; ADAD2; LOC654780; KCNG4*
20	rs2297122	1.03E-11	854807	T / C	0.04 / 0.09	0.33 (0.24 - 0.46)	*FAM110A; ANGPT4*
9	rs79128444	1.73E-11	31258879	G / A	0.07 / 0.15	0.44 (0.35 - 0.56)	-
5	rs60507859	6.33E-11	57362947	A / C	0.03 / 0.10	0.35 (0.25 - 0.48)	*LINC02101*
3	rs111529683	6.01E-10	155001404	A / G	0.05 / 0.10	0.40 (0.30 - 0.53)	*LINC01487; STRIT1*
2	rs3820864	1.97E-09	169012074	A / C	0.40 / 0.29	1.58 (1.36 - 1.84)	*STK39*
6	rs241410	2.81E-09	32861651	G / C	0.54 / 0.43	1.55 (1.34 - 1.79)	*PSMB8; PSMB8-AS1; TAP1; PSMB9; LOC100294145; HLA-DMB*
6	rs3131634	3.07E-09	31477189	A / G	0.03 / 0.08	0.34 (0.24 - 0.49)	*HCP5; HCG26; MICB-DT; MICB; MCCD1; ATP6V1G2-DDX39B; DDX39B; SNORD117; SNORD84; DDX39B-AS1; ATP6V1G2; NFKBIL1*
11	rs4520594	5.45E-09	11540861	A / G	0.12 / 0.06	2.33 (1.75 - 3.09)	*GALNT18*
6	rs6938961	7.72E-09	28280297	A / G	0.19 / 0.11	1.89 (1.53 - 2.35)	*ZSCAN26; PGBD1; ZSCAN31; ZKSCAN3*
6	rs9258051	1.26E-08	29665228	A / G	0.58 / 0.47	1.55 (1.33 - 1.81)	*MOG; ZFP57; HLA-F; HLA-F-AS1*
15	rs146875672	1.27E-08	44775198	T / C	0.04 / 0.09	0.41 (0.30 - 0.55)	*CTDSPL2*
12	rs4091651	1.88E-08	96880756	G / A	0.33 / 0.25	1.60 (1.36 - 1.88)	*CFAP54*
2	rs1549343	2.05E-08	171093785	G / A	0.28 / 0.19	1.63 (1.37 - 1.93)	*MYO3B*
6	rs241412	2.18E-08	32861616	T / G	0.42 / 0.32	1.55 (1.33 - 1.81)	*PSMB8; PSMB8-AS1; TAP1; PSMB9; LOC100294145; HLA-DMB*
16	rs77286017	2.45E-08	70738119	A / G	0.04 / 0.09	0.44 (0.33 - 0.58)	*IL34; MTSS2; VAC14*
12	rs11170867	3.44E-08	54717410	C / T	0.17 / 0.11	1.85 (1.48 - 2.29)	*CBX5; SCAT2; HNRNPA1; NFE2; COPZ1; MIR148B; GPR84-AS1; GPR84; ZNF385A*
13	rs2057412	3.99E-08	49982570	C / T	0.08 / 0.15	0.54 (0.43 - 0.67)	*CAB39L; SETDB2; SETDB2-PHF11*
1	rs6426281	4.97E-08	246310557	G / A	0.07 / 0.14	0.50 (0.39 - 0.64)	*SMYD3*

Abbreviations. CHR: Chromosome; SNP: Single Nucleotide Polymorphism; P value: adjusted logistic regression p-value; BP: Base pair, Genome Reference Consortium Human Build 37; MAF: minor allele frequency in case/control; OR: adjusted odds ratio, minor allele as reference; CI: confidence interval, lower 95% and upper 95% confidence interval. Gene annotations as per NCBI RefSeq track.

doi: 10.18176/jiaci.0978