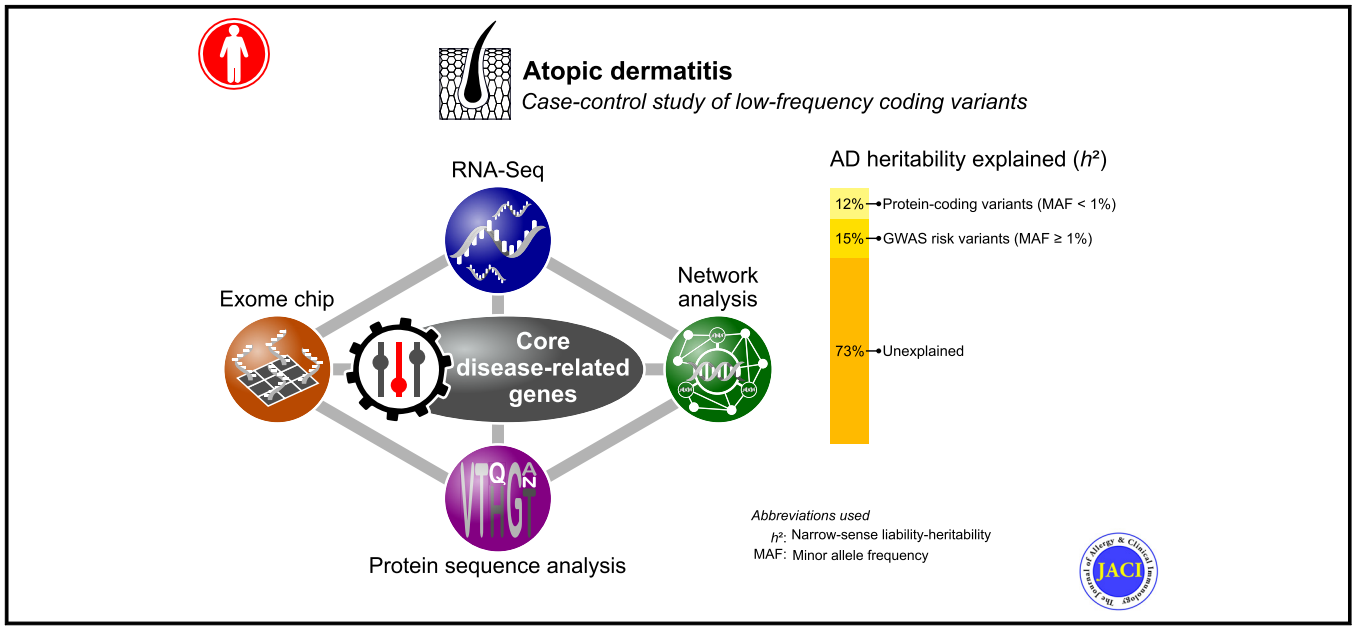
Protein-coding variants contribute to the risk of atopic dermatitis and skin-specific gene expression

Sören Mucha, MSc,[a]* Hansjörg Baurecht, PhD,[b,c]* Natalija Novak, MD,[d] Elke Rodríguez, PhD,[b] Saptarshi Bej, MSc,[e] Gabriele Mayr, PhD,[a] Hila Emmert, PhD,[b] Dora Stölzl, MD,[b] Sascha Gerdes, MD,[b] Eun Suk Jung, MD,[a,f] Frauke Degenhardt, MSc,[a] Matthias Hübenthal, PhD,[a,b] Eva Ellinghaus, PhD,[a] Jan Christian Kässens, PhD,[a] Lars Wienbrandt, PhD,[a] Wolfgang Lieb, MD,[g] Martina Müller-Nurasyid, PhD,[h,i,j] Melanie Hotze, PhD,[b] Nick Dand, PhD,[k] Sarah Grosche, PhD,[l,m] Ingo Marenholz, MD,[l,m] Andreas Arnold, MD,[n] Georg Homuth, PhD,[o] Carsten O. Schmidt, MD,[p] Ulrike Wehkamp, MD,[b] Markus M. Nöthen, MD,[q] Per Hoffmann, PhD,[q] Lavinia Paternoster, PhD,[r] Marie Standl, PhD,[s] on behalf of the Early Genetics and Lifecourse Epidemiology (EAGLE) Eczema Consortium, Klaus Bønnelykke, MD,[t] Tarunveer S. Ahluwalia, PhD,[t,u] Hans Bisgaard, MD,[t] Annette Peters, PhD,[s] Christian Gieger, PhD,[v] Melanie Waldenberger, PhD,[v] Holger Schulz, PhD,[s] Konstantin Strauch, PhD,[h,i] Thomas Werfel, MD,[u,w] Young-Ae Lee, PhD,[j,k] Markus Wolfien, MSc,[e] Philip Rosenstiel, MD,[a] Olaf Wolkenhauer, PhD,[e] Stefan Schreiber, MD,[a,x] Andre Franke, PhD,[a]‡ Stephan Weidinger, MD,[b]‡ and David Ellinghaus, PhD[a]‡ *Kiel, Regensburg, Bonn, Rostock, Neuherberg, Munich, Berlin, Greifswald, and Hannover, Germany; Seoul, Korea; London and Bristol, United Kingdom; and Gentofte, Denmark*

GRAPHICAL ABSTRACT



From [a]the Institute of Clinical Molecular Biology and [s]the Institute of Epidemiology and Biobank PopGen, Christian Albrechts University of Kiel; [b]the Department of Dermatology, Venereology and Allergy, University Hospital Schleswig-Holstein, Campus Kiel; [c]the Department for Epidemiology and Preventive Medicine, University of Regensburg; [d]the Department of Dermatology and Allergy, University Hospital Bonn; [e]the Department of Systems Biology and Bioinformatics, University of Rostock; [f]the Department of Internal Medicine and Institute of Gastroenterology, Yonsei University College of Medicine, Seoul; [h]the Institute of Genetic Epidemiology, [s]the Institute of Epidemiology, and [v]the Research Unit of Molecular Epidemiology and Institute of Epidemiology II, Helmholtz Zentrum München–German Research Center for Environmental Health, Neuherberg; [i]the Chair of Genetic Epidemiology, IBE, Faculty of Medicine, Ludwig-Maximilians-University Munich; [j]the Department of Internal Medicine I (Cardiology), Hospital of the Ludwig-Maximilians-University Munich; [k]the School of Basic & Medical Biosciences, Faculty of Life Sciences & Medicine, King's College London; [l]Pediatric Allergology, Experimental and Clinical Research Center, Charité Universitätsmedizin Berlin; [m]Max-Delbrück-Centrum (MDC) for Molecular Medicine, Berlin; [n]the Clinic and Polyclinic of Dermatology and [p]the Institute for Community Medicine, Study of Health in Pomerania/KEF, University Medicine Greifswald; [o]the Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz-Arndt-University Greifswald; [q]the Institute of Human Genetics, University of Bonn; [r]Medical Research Council (MRC) Integrative Epidemiology Unit, Bristol Medical School, and the School of Social and Community Medicine, University of Bristol; [t]Copenhagen Prospective Studies on Asthma in Childhood (COPSAC), Herlev and Gentofte Hospital, Gentofte; [u]Steno Diabetes Center Copenhagen, Gentofte; [w]Department of Dermatology and Allergy, Division of Immunodermatology and Allergy Research, Hannover Medical School; and [x]First Medical Department, University Hospital Schleswig-Holstein, Kiel.

*These authors contributed equally to this work.

‡These authors jointly supervised this work.

Background: Fifteen percent of atopic dermatitis (AD) liability-scale heritability could be attributed to 31 susceptibility loci identified by using genome-wide association studies, with only 3 of them (*IL13*, IL-6 receptor *[IL6R]*, and filaggrin *[FLG]*) resolved to protein-coding variants.

Objective: We examined whether a significant portion of unexplained AD heritability is further explained by low-frequency and rare variants in the gene-coding sequence.

Methods: We evaluated common, low-frequency, and rare protein-coding variants using exome chip and replication genotype data of 15,574 patients and 377,839 control subjects combined with whole-transcriptome data on lesional, nonlesional, and healthy skin samples of 27 patients and 38 control subjects.

Results: An additional 12.56% (SE, 0.74%) of AD heritability is explained by rare protein-coding variation. We identified docking protein 2 *(DOK2)* and CD200 receptor 1 *(CD200R1)* as novel genome-wide significant susceptibility genes. Rare coding variants associated with AD are further enriched in 5 genes (IL-4 receptor *[IL4R]*, *IL13*, Janus kinase 1 *[JAK1]*, *JAK2*, and tyrosine kinase 2 *[TYK2]*) of the *IL13* pathway, all of which are targets for novel systemic AD therapeutics. Multiomics-based network and RNA sequencing analysis revealed *DOK2* as a central hub interacting with, among others, *CD200R1*, *IL6R*, and signal transducer and activator of transcription 3 *(STAT3)*. Multitissue gene expression profile analysis for 53 tissue types from the Genotype-Tissue Expression project showed that disease-associated protein-coding variants exert their greatest effect in skin tissues.

Conclusion: Our discoveries highlight a major role of rare coding variants in AD acting independently of common variants. Further extensive functional studies are required to detect all potential causal variants and to specify the contribution of the novel susceptibility genes *DOK2* and *CD200R1* to overall disease susceptibility. (J Allergy Clin Immunol 2020;145:1208-18.)

Key words: *Atopic dermatitis, exome chip association analysis, network analysis, protein sequence and structural domain analysis, RNA sequencing*

Abbreviations used

AD:	Atopic dermatitis
CD200R1:	CD200 receptor 1 gene
DOK2:	Docking protein 2 gene
EAGLE:	Early Genetics and Lifecourse Epidemiology
GTEx:	Genotype-Tissue Expression
GWAS:	Genome-wide association study
IL6R:	IL-6 receptor gene
LD:	Linkage disequilibrium
MAF:	Minor allele frequency
MS:	Multiple sclerosis
OR:	Odds ratio
RasGAP:	Ras GTPase activating protein
RNA-seq:	RNA sequencing
SNP:	Single nucleotide polymorphism

Atopic dermatitis (AD; Mendelian Inheritance in Man: 603165) is the most common chronic inflammatory skin disorder, affecting 15% to 20% of children and 5% to 10% of adults (approximately 280 million persons worldwide), and is the leading cause of the nonfatal disease burden conferred by skin conditions.[1] Given its high genetic heritability (71% to 84% in Europeans),[2] finding causal genes is a crucial step for developing effective preventive and therapeutic approaches for AD. Thus far, genome-wide association studies (GWASs) have identified 31 specific genomic regions associated with AD susceptibility.[3-9] The reported susceptibility variants are common (n = 31 with minor allele frequencies [MAFs] ≥ 5%) and mostly located in noncoding DNA regions of the genome, have rather small effect sizes (odds ratio [OR] < 1.15), and have a largely unclear functional significance.[9] Notable exceptions are low-frequency null mutations in the gene encoding the epidermal structural protein filaggrin *(FLG)*, which lead to a reduction in biologically active filaggrin peptides and a complex perturbation of skin barrier function,[10] as well as common missense variants in genes encoding the T_H2 signature cytokine IL-13 (*IL13*; rs20541),[11] as well as the IL-6 receptor (*IL6R*; rs2228145).[12] Recently, an exome chip–based association study of low-frequency variation across all autosomal exons in multiple sclerosis (MS) cohorts of European ancestry led to the detection of low-frequency MS-associated coding variants for 4 genes that were missed by

This work was supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (SysInflame grant 01ZX1606A and GB-XMAP grant 01ZX1709), as well as the German ELIXIR node de.NBI (de.STAIR grant 031L0106C). The project was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC 2167-390884018). The project received infrastructure support from the DFG Excellence Cluster no. 306 "Inflammation at Interfaces" and the PopGen Biobank (Kiel, Germany). The Kooperative Gesundheitsforschung in der Region Augsburg (KORA) research platform (Cooperative Research in the Region of Augsburg) was initiated and financed by the Helmholtz Zentrum München–German Research Center for Environmental Health, which is funded by BMBF and the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. SHIP (Study of Health in Pomerania) is part of the Community Medicine Research net (CMR) of the University of Greifswald, Greifswald, Germany, which is funded by BMBF (grants 01ZZ9603, 01ZZ0103, and 01ZZ0403) and the Ministry of Cultural Affairs, as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network "Greifswald Approach to Individualized Medicine (GANI_MED)" funded by BMBF (grant 03IS2061A). Exome chip data have been supported by BMBF (grant no. 03Z1CN22) and the Federal State of Mecklenburg-West Pomerania. L.P. was supported by a UK Medical Research Council fellowship (MR/J012165/1). This work was supported by BIOMAP (Biomarkers in Atopic Dermatitis and Psoriasis), a project funded by the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement no. 821511 and in-kind contributions of the participating pharma companies. The Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program and EFPIA.

Disclosure of potential conflict of interest: The authors declare that they have no relevant conflicts of interest.

Received for publication June 24, 2019; revised September 9, 2019; accepted for publication October 7, 2019.

Available online November 9, 2019.

Corresponding authors: Stephan Weidinger, MD, Arnold-Heller-Straße 3, Haus 19, 24105 Kiel, Germany. E-mail: sweidinger@dermatology.uni-kiel.de. OR: David Ellinghaus, PhD, Rosalind-Franklin-Str 12, 24105 Kiel, Germany. E-mail: d.ellinghaus@ikmb.uni-kiel.de.

The CrossMark symbol notifies online readers when updates have been made to the article such as errata or minor corrections

0091-6749/$36.00

https://doi.org/10.1016/j.jaci.2019.10.030

using a previous large-scale MS consortium GWAS[13] and that explain another 5% of MS heritability, thus reopening the debate on the contribution of low-frequency and rare variants to disease risk in complex diseases.

To systematically evaluate the contribution of genetic variation to the genetic architecture of AD on an exome-wide scale, particularly protein-altering variants of low or rare frequency, we profiled 1,913 patients with AD and 14,295 control subjects in 2 German cohorts using the Illumina HumanExome BeadChip (exome chip, see Table E1 in this article's Online Repository at www.jacionline.org). The exome chip captures approximately 88% of low-frequency and rare coding variants (nonsynonymous, splice-site, and stop-altering variants; MAFs between 0.01% and 5%) present in Europeans.[14] Suggestive significant novel associations ($P_{\text{exome chip}} < 1 \times 10^{-5}$) were taken forward to replication genotyping in a third German cohort of 1,789 AD cases and 3,272 control subjects, a Danish exome chip case-control study of 292 severe AD cases and 650 control subjects, GWAS association statistics of the Early Genetics and Lifecourse Epidemiology (EAGLE) eczema consortium of 2,298 independent AD cases and 7,802 control subjects (see Table E1), and association summary statistics of 361,132 subjects from the UK Biobank with self-reported information (see Table E1). Gene-based tests and Bayesian fine-mapping analysis, as well as whole-transcriptome RNA sequencing (RNA-seq), immunohistochemistry, and variant protein analyses, were conducted to elucidate potential functional consequences of coding variation associated with AD. Finally, we performed multiomics-based network, pathway gene set, and gene expression tissue profile analyses and quantified the overall contribution of exome chip variation to AD risk.

METHODS

The Methods section in this article's Online Repository at www.jacionline.org provides details on the methods used in this study.

Study samples and genotyping

All cases received a diagnosis of AD by a dermatologist except for UK Biobank cases. All participants provided written informed consent, and the study was approved by the ethics boards of the participating institutions in agreement with the Declaration of Helsinki principles.

German discovery cohort 1. German patients with AD (n = 1,056) were recruited at the Department of Dermatology at Christian Albrechts University Kiel, the Department of Dermatology and Allergy at the Technical University of Munich, and the Department of Dermatology and Allergy at the University of Bonn. Data from healthy control subjects (n = 7,026) were obtained from the PopGen Biobank,[15] the Kooperative Gesundheitsforschung in der Region Augsburg (KORA) S4 survey (an independent population-based sample from the general population living in the region of Augsburg, southern Germany),[16] and the Heinz-Nixdorf Recall cohort (Bonn, Germany).[17] Both patients with AD and control subjects were genotyped with Illumina HumanExome-12 v1.0 BeadChips (see Table E1 and Fig E1 in this article's Online Repository at www.jacionline.org).

German discovery cohort 2. German patients with AD (n = 1,051) were recruited from dermatology clinics in Kiel or Hannover (the University of Kiel and Medizinische Hochschule of Hannover). The AD cases were genotyped by using Illumina HumanCoreExome-24 v1.0 A or HumanCoreExome-24 v1.1 A BeadChips. Data from healthy control subjects (n = 8,135) were obtained from the Study of Health in Pomerania (SHIP) and SHIP-TREND cohorts (from SHIP, a prospective longitudinal population-based cohort study in West Pomerania).[18] All German control subjects were genotyped with Illumina HumanExome-12 v1.0 BeadChips (see Table E1 and Fig E1).

German replication cohort. German patients with AD (n = 1,789) were recruited from dermatology clinics in Kiel or Berlin (University Children's Hospital, Charité Universitätsmedizin Berlin, as part of the Genetic Studies in Nuclear Families with AD study). Data from healthy control subjects (n = 3,272) were obtained from University Hospital in Kiel and Lübeck at the Institute of Transfusional Medicine (see Table E1).

Danish replication cohort. All Danish patients with AD (n = 292) are hospitalized severe cases from the COPSAC eczema registry. Healthy control subjects (n = 650) were obtained from the COPSAC2000 and COPSAC2010 birth cohorts[19] in Copenhagen, Denmark. Both cases and control subjects were genotyped on the Illumina Infinium OmniExpressExome-8 v1.4 BeadChip (see Table E1).

EAGLE GWAS replication cohorts. We used imputed summary statistics of the EAGLE Eczema Consortium for the discovery cohorts (excluding 23andMe; https://data.bris.ac.uk/data/dataset/28uchsdpmub118uex26ylacqm) comprising 11,294,660 single nucleotide polymorphism (SNP) markers with MAFs of 1% or greater and 10,788 AD cases and 30,047 control subjects (see Table E1).[9] For replication analysis, we used only European studies independent from our German discovery cohorts and in which AD diagnosis was ascertained by a dermatologist.

UK Biobank. Because only 33 (primary diagnosis; field 41203) and 50 (secondary diagnosis; field 41204) patients have been given a diagnosis of "atopic dermatitis" (International Classification of Diseases, Tenth Edition, code L30) in UK Biobank, we used questionnaire information from UK Biobank (March 2018 release). The key words "atopic dermatitis," "eczema," or both are contained in the self-reported data fields "Non-cancer illness code; self-reported: eczema/dermatitis" (data field 20002), "Age hay fever, rhinitis or eczema diagnosed" (data field 3761), and 3 others (data fields 6152, 41202, and 41203). For the 3 lead variants (*DOK2*: rs34215892 and rs56094005; *CD200R1*: rs9865242), we downloaded imputed summary association statistics (http://www.nealelab.is/uk-biobank/; http://www.nealelab.is/uk-biobank/faq; release March 2018; see Table E1) for (1) 83,407 cases with eczema, allergic rhinitis, and/or hay fever (data field 20002) versus 277,120 control subjects and (2) 9,312 cases with eczema or dermatitis (data field 3761) versus 351,820 control subjects.

An overview of the study design is shown in Fig E2 in this article's Online Repository at www.jacionline.org, and detailed characteristics of the discovery and replication case-control cohorts are provided in Table E1.

RESULTS

Exome chip single-variant association analysis

In our exome chip discovery search for common, low-frequency, and rare variant associations, 2 German AD discovery cohorts were combined through a meta-analysis, resulting in single variant association analysis score statistics of 143,884 genotyped and 1,357,289 single nucleotide variants after imputation with information scores of 0.5 or greater (see the Methods section, Figs E2, E3, and E5 and Table E2 in this article's Online Repository at www.jacionline.org). Four hundred thirty-eight and 1,331 SNPs within 7 and 25 loci were identified with genome-wide significance ($P_{\text{exome chip}} < 5 \times 10^{-8}$) and suggestive association ($P_{\text{exome chip}} < 1 \times 10^{-5}$), respectively (see Fig E4 and Table E3 in this article's Online Repository at www.jacionline.org).

To identify novel susceptibility variants outside of established AD GWAS loci (see Table E4 in this article's Online Repository at www.jacionline.org), 11 suggestively associated SNPs ($P_{\text{exome chip}} < 10^{-5}$ with MAF $\geq$ 1%) were selected based on the linkage disequilibrium (LD) clumping method (see the Methods section and Fig E7 in this article's Online Repository at www.jacionline.org) and carried forward for replication (see Table E5

TABLE I. Exome chip discovery and replication genotyping, as well as UK Biobank replication (self-reported AD) single-marker association analysis revealed *DOK2* and *CD200R1* as genome-wide significant AD susceptibility genes

dbSNP ID AA substitution	Position (chr:bp)	Gene	Variant type	A1/A2	AF_{cases}	$AF_{control\ subjects}$	Discovery, n_{cases}/ $n_{control\ subjects}$	Discovery *P* value/OR (95% CI)	Replication, n_{cases}/ $n_{control\ subjects}$	Replication *P* value/OR (95% CI)	Combined *P* value/OR (95% CI)
rs34215892 P274L	8:21767240	*DOK2*	Missense	A/G	0.025	0.039	1,913/14,295	9.83×10^{-7}; 0.61 (0.34–0.86)	3,962/10,076	1.35×10^{-5}; 0.68 (0.56-0.79)	**2.15×10^{-10}; 0.64 (0.53–0.76)**
rs9865242 E312Q	3:112642568	*CD200R1*	Missense	C/G	0.457	0.435	1,913/14,295	5.80×10^{-6}; 1.18 (1.15–1.22)	4,349/11,724	3.38×10^{-3}; 1.14 (1.09–1.19)	1.17×10^{-7}; 1.16 (1.13–1.19)

					UK Biobank summary statistics for AD, allergic rhinitis, and/or hay fever (self-reported)				UK Biobank summary statistics for AD or dermatitis (self-reported)			
dbSNP ID AA substitution	Position (chr:bp)	Gene	Variant type	A1/A2	AF_{cases}	$AF_{control\ subjects}$	Phenotype 1, n_{cases}/ $n_{control\ subjects}$	*P* value	AF_{cases}	$AF_{control\ subjects}$	Phenotype 2, n_{cases}/ $n_{control\ subjects}$	*P* value
rs34215892 P274L	8:21767240	*DOK2*	Missense	A/G	0.027	0.029	83,407/277,120	3.35×10^{-6}	0.025	0.029	9,312/351,820	2.50×10^{-3}
rs9865242 E312Q	3:112642568	*CD200R1*	Missense	C/G	0.474	0.467	83,407/277,120	**1.35×10^{-8}**	0.487	0.468	9,312/351,820	8.33×10^{-8}

Because exome chip discovery and replication genotyping studies comprised only patients with AD ascertained by a dermatologist in comparison with the UK Biobank replication study with only self-reported AD, association results were listed separately. *P* values and ORs were calculated with respect to the minor allele, and genome-wide significant *P* values ($P < 5 \times 10^{-8}$) are indicated in boldface. All patients with AD from the discovery and replication panels had received a diagnosis of AD (eczema) by a dermatologist. rs34215892 and rs9865242 were genotyped (nonimputed) in discovery and replication panels (except for the EAGLE GWAS replication data). Association results from the UK Biobank (self-reported phenotypes; see Table E1) are listed separately.

A1, minor allele; *A2*, major allele; *AA*, Amino acid; *AF*, allele frequency of A1; *bp*, genomic position from National Center for Biotechnology Information dbSNP build v150 (genome build hg19); *chr*, chromosome of the marker; *Gene*, candidate gene; *OR*, estimated odds ratio.

in this article's Online Repository at www.jacionline.org). Using the genome-wide threshold of 5×10^{-8} for the combined analysis of discovery and replication (see the Methods section in this article's Online Repository), we identified a novel low-frequency missense variant in exon 3 of the docking protein 2 gene (*DOK2* at 8p21.3; rs34215892 [p.P274L]; $P_{exome\ chip} = 9.83 \times 10^{-7}$; genotyped variant), which consistently and robustly replicated in 3 independent cohorts ($P_{German} = 3.75 \times 10^{-4}$, $P_{Denmark} = 7.60 \times 10^{-3}$, $P_{EAGLE} = 4.35 \times 10^{-2}$, $P_{combined} = 2.15 \times 10^{-10}$, and $OR_{combined} = 0.64$; Table I and see Table E5). Furthermore, we detected a novel association between AD and a common missense variant at 3q13.2 (rs9865242; p.E312Q; $P_{combined} = 1.17 \times 10^{-7}$; $OR_{combined} = 1.16$; genotyped variant; see Table E5) located in exon 7 of CD200 receptor 1 *(CD200R1)* and 266.51 kb upstream of the previously reported intergenic locus 3q13.2 (*CCDC80*, rs12634229) described only in a Japanese AD cohort[6] thus far ($r^2_{rs9865242\text{-}rs12634229} = 0.005$). A lookup of association results from UK Biobank for the self-reported broad allergic disease phenotype "AD (eczema), allergic rhinitis and/or hayfever" (see the Methods section in this article's Online Repository)[20] further confirmed association signals for *DOK2* (rs34215892; $P_{UK\text{-}Biobank} = 3.35 \times 10^{-6}$) and *CD200R1* (rs9865242; $P_{UK\text{-}Biobank} = 1.35 \times 10^{-8}$; Table I).

Exome chip gene-based association analysis

Because of insufficient statistical power to perform single-marker tests for rare variants, we performed gene-based association analysis, in which we evaluated the cumulative effects of low-frequency and rare variants (nonsynonymous, stop-gain, and essential splice-site variants; n = 118,816 exome chip variants, excluding imputed variants) for each gene from autosomes (see the Methods section in this article's Online Repository). Only *DOK2* met the exome-wide significance threshold (see the Methods section in this article's Online Repository) and exhibited a stronger association signal than compared with single variant analysis ($P_{combined\text{-}DOK2} = 4.23 \times 10^{-13}$; Table II and see Table E6 in this article's Online Repository at www.jacionline.org) comprising 12 protein-altering variants of which 2 were of low frequency and LD independent (rs34215892 and rs56094005; $1\% \leq MAF < 5\%$) and 10 were rare ($MAF < 1\%$, Fig 1). We genotyped the second low-frequency variant rs56094005 (p.L138S; $P_{rs56094005\text{-}exomechip} = 4.31 \times 10^{-3}$; genotyped variant) in our replication set in addition to rs34215892 and successfully confirmed the single SNP ($P_{rs56094005\text{-}replication} = 5.96 \times 10^{-3}$, see Table E5) and the aggregated *DOK2* association signal ($P_{DOK2\text{-}replication} = 1.54 \times 10^{-6}$; see Table E6).

Bayesian fine mapping and functional annotation

As a next step toward understanding the functional causality of the identified AD-associated variants, we carried out Bayesian fine-mapping and functional annotation analysis for *CD200R1* and *DOK2* loci using imputed exome chip data of our discovery cohorts (see the Methods section in this article's Online Repository). Fine mapping strengthened our hypothesis that the lead variants rs9865242 *(CD200R1)* and rs34215892 *(DOK2)* are most likely causal (in the context of fine mapping), with posterior probabilities of 97.2% and 44.3%, respectively (see Fig E6 in this article's Online Repository at www.jacionline.org). rs34215892 overlaps enhancer histone marks and DNase peaks in 15 and 12 different tissues, respectively, in each case, including the skin, and is predicted to affect protein-binding and regulatory motifs (see Table E7 in this article's Online Repository at www.jacionline.org). The second *DOK2* lead variant, rs56094005 (Fig 1), overlaps promoter histone marks and DNase peaks in 10 tissues, in each case including T and B cells, and is in perfect LD ($r^2 = 1$) with the intronic variant rs118162691 ($P_{rs118162691\text{-}exomechip} = 4.17 \times 10^{-3}$,

TABLE II. Meta-analysis of gene-based aggregation tests for *DOK2* increased the genome-wide significant association signal by more than 2 orders of magnitude in comparison with single-marker analysis (Table I), indicating that multiple rare variants (with 10 of 12 variants predicted to be pathogenic; Fig 1) contribute to the association signal

Gene	dbSNP ID	Position (chr:bp)	AF_{mean}	AA	Prediction	Discovery, $n_{cases}/n_{control\ subjects}$ $pSKAT_{discovery}$	Replication, $n_{cases}/n_{control\ subjects}$ $pSKAT_{replication}$	Combined, $n_{cases}/n_{control\ subjects}$ $pSKAT_{combined}$
DOK2	rs2242241†	8:21766881	0.00028	S394A	Tolerated, possibly damaging	1,913/14,295 **2.61×10^{-7}**	3,932/10,076 **1.54×10^{-6}**	5,845/24,371 **4.23×10^{-13}**
	rs145725971*	8:21767033	0.00065	H343R	Tolerated, possibly damaging			
	rs145405180†	8:21767148	0.00139	A305T	Tolerated, benign			
	rs34215892†	8:21767240	0.03708	P274L	Damaging, probably damaging			
	rs74909419*	8:21767265	0.00019	R266W	Damaging, probably damaging			
	rs141482665*	8:21767322	0.00012	N247D	Tolerated, probably damaging			
	rs200168233*	8:21767414	0.00015	R216H	Damaging, probably damaging			
	rs200503110†	8:21767417	0.00019	R215H	Damaging, probably damaging			
	rs149080191*	8:21768370	0.00009	ESS				
	rs56094005†	8:21769432	0.04026	L138S	Damaging, probably damaging			
	rs201025320†	8:21770011	0.00003	R25H	Damaging, probably damaging			
	rs142660088†	8:21771080	0.00043	L11F	Damaging, probably damaging			

Bonferroni-corrected exome chip significant gene-based *P* values ($P_{gene} < .05/15{,}998 = 3 \times 10^{-6}$; 15,998 genes) are indicated in boldface. Cohort-specific association details are shown in Table E6. All 12 variants were genotyped (nonimputed) in the discovery and replication panels (except for the EAGLE GWAS replication data).

AA, Amino acid substitution; AF_{mean}, mean allele frequency of minor allele from discovery panels (see Table E5); *bp*, genomic position from National Center for Biotechnology Information dbSNP build v150 (genome build hg19); *chr*, chromosome of the marker; *ESS*, essential splice site; *Gene*, candidate gene; *Prediction*, SIFT prediction or PolyPhen-2 prediction; *pSKAT*, *P* value of the sequence kernel association test.

*Risk variant (in context of the OR).

†Protective variant.

MAF = 3.9%), which overlaps enhancer histone marks in 9 different tissues, including the skin (see Table E7). The *CD200R1* missense variant rs9865242 has been suggested as a cis–expression quantitative trait locus for *GTPBP8* in whole blood.[21] Ten variants in perfect LD with rs9865242 and further overlapping enhancer histone marks in blood predicted to alter regulatory motifs are located in a conserved region or are found as expression quantitative trait loci for *CD200R1* (see Table E7).[22]

Immunohistochemistry and whole-transcriptome mRNA-seq data analysis

Immunohistochemistry was used to determine the location of DOK2 in skin tissue (see the Methods section in this article's Online Repository). It showed strong epidermal staining with clear differences among AD lesional, AD nonlesional, and healthy skin (see Fig E8 in this article's Online Repository at www.jacionline.org). DOK2 is predominantly expressed in lymphocytes, and the increased abundance of DOK2 in lesional AD skin correlates with the degree of lymphocyte infiltration (see Fig E8). Moreover, we observed significantly increased *DOK2* and *CD200R1* mRNA expression in whole-transcriptome mRNA-seq data (see the Methods section in this article's Online Repository) on lesional compared with nonlesional skin samples of 27 patients with AD ($P_{DOK2} = 4.2 \times 10^{-5}$, $P_{CD200R1} = 2.2 \times 10^{-5}$) and compared with skin from 38 healthy subjects ($P_{DOK2} = 8.8 \times 10^{-11}$, $P_{CD200R1} = 2.2 \times 10^{-7}$), as well as in AD nonlesional skin compared with healthy skin ($P_{DOK2} = 4.5 \times 10^{-3}$, $P_{CD200R1} = 2.0 \times 10^{-2}$; see Fig E9, *A* and *B*, in this article's Online Repository at www.jacionline.org). We also observed a trend toward increased expression of *DOK1*, the protein of which is a heterodimeric partner for DOK2, in lesional skin compared with nonlesional ($P_{DOK1} = 1.1 \times 10^{-1}$) or healthy ($P_{DOK1} = 3.1 \times 10^{-4}$) skin, as well as nonlesional skin compared with healthy skin ($P_{DOK1} = 4.0 \times 10^{-3}$; see Fig E9, *C* and *D*).

In silico variant protein analysis

To construct a first hypothetical model of whether *CD200R1* and *DOK2* missense lead variants are likely to interfere with functionally active domains on the protein level, we performed an extensive literature search and further conducted protein domain analyses of DOK2 and the CD200/CD200R1 receptor complex (Fig 2 and see the Methods section in this article's Online Repository).[23-25] Members of the DOK adapter protein family act as regulators of cell-stimulatory signals by serving as substrates for diverse receptor and cytoplasmic kinases, and the highly similar and interacting members DOK1 and DOK2 are involved in downregulation of immune receptor signaling in $CD4^+$ T cells, as well as myeloid cells, such as macrophages and neutrophils.[26] It is assumed that the inhibitory role of DOK2 is accomplished by recruiting and activating Ras GTPase-activating protein (RasGAP) to inhibit Ras and thus to suppress proinflammatory extracellular signal-regulated kinase, c-Jun N-terminal kinase, and mitogen-activated protein kinase pathways for the DOK2 response to CD200R1 in human myeloid cells.[23] Activated RasGAP inhibits mitogen-activated protein kinase signaling and subsequently reduces production of proinflammatory cytokines, such as TNF-α, INF-γ, IL-1, IL-17, IL-6, and IL-8.[27,28] Using our protein sequence and structure analyses, we observed that variant rs56094005 (p.L138S) locates within a linker sequence between the Pleckstrin homology domain and the phosphotyrosine-binding domain adjacent to the

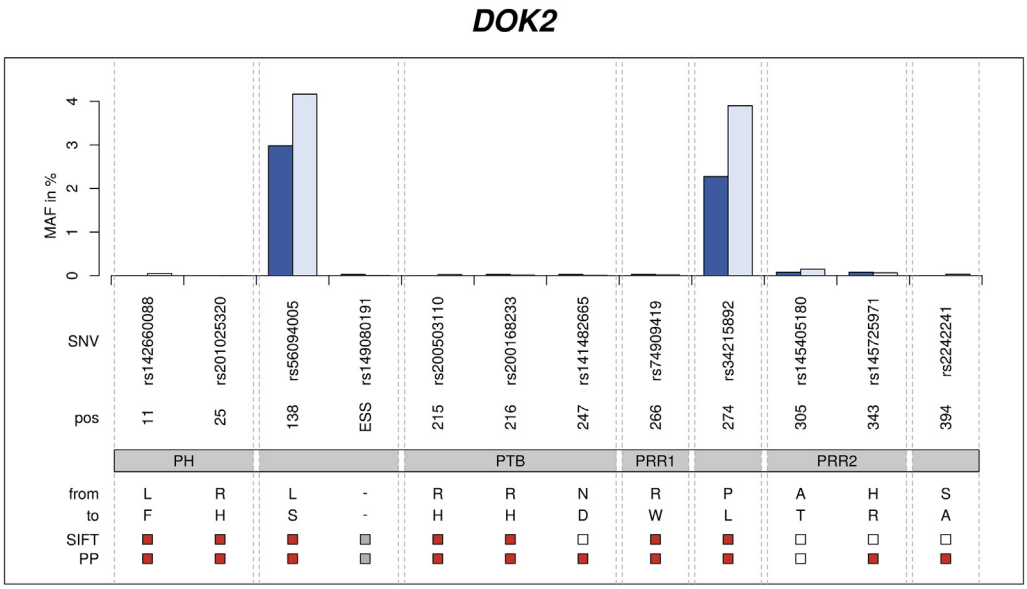


FIG 1. Exome chip association analysis identified 2 low-frequency (rs34215892 and rs56094005; 1% ≤ MAF < 5%) and 10 rare (MAF < 1%) coding variants contributing to the genome-wide significant association signal at *DOK2* ($P_{DOK2} = 4.23 \times 10^{-13}$, Table II). MAFs in patients with AD *(dark blue columns)* and control subjects *(light blue columns)* are shown from discovery exome chip meta-analysis (see Table E1). Ten of 12 variants were predicted to be pathogenic (see the Methods section in this article's Online Repository), suggesting that multiple missense variants contribute to the gene-based association signal. Both low-frequency lead variants are LD independent ($r^2_{rs56094005\text{-}rs34215892} = 0.0015$). Variant effect predictions (by SIFT and PolyPhen-2 *[PP]*) depicted in red (or white or gray, respectively) represent amino acid substitutions predicted to be potentially damaging (or probably not damaging or prediction not possible, respectively). *A*, Alanine; *D*, aspartic acid; *ESS*, essential splice site; *F*, phenylalanine; *H*, histidine; *L*, leucine; *N*, asparagine; *P*, proline; *PH*, Pleckstrin homology domain; *pos*, amino acid position; *PRR1*, proline-rich region 1; *PRR2*, proline-rich region 2; *PTB*, phosphotyrosine-binding domain; *R*, arginine; *S*, serine; *SNV*, single nucleotide variant; *T*, threonine; *W*, tryptophan.

Y139 phosphorylation-dependent DOK1 interaction site and might interfere with heterodimerization of DOK1 and DOK2, which is required for full phosphorylation of the 2 proteins and signaling.[29] Variant rs34215892 (p.P274L), which is located in the invariant RasGAP-SH2 binding consensus motif YxxPxD,[29] likely affects local protein structure conformation because of the unique structural rigidity of the proline side chain and therefore is predicted to interfere with RasGAP signaling. Assuming an important stabilizing structural role of the proline within the binding motif, the variant p.P274L is likely to disturb the RasGAP activation by DOK2. The amino acid substitution of E to Q of variant rs9865242 (p.E312Q) causes a loss of negative charge in proximity of the NPLY motif and protein interaction site and might therefore modify protein-protein contacts and signaling.[24]

AD core network construction and differential gene expression analysis of core genes

To assess possible interactions of the genes *CD200R1* and *DOK2* with candidate genes from AD GWAS loci,[3-9] we generated an AD core interaction network using network prioritization algorithms, which make use of protein-protein, protein-gene, coexpression, and shared protein domain data from public repositories (see the Methods section, Fig E10, and Table E8 in this article's Online Repository at www.jacionline.org). We identified *DOK2* as a central hub node interacting with *CD200R1* and the functionally established AD susceptibility genes signal transducer and activator of transcription 3 *(STAT3)*,[9] *MICB*,[30] *CLEC16A*,[8] and *IL6R*[12] (Fig 3, *A*). We used our aforementioned whole-transcriptome mRNA-seq data (see the Methods section in this article's Online Repository) to assess whether expression levels from genes of our AD core network are differentially expressed in lesional, nonlesional, and healthy skin from patients with AD and healthy subjects. Twenty-two of the 30 AD core genes (Fig 3, *A*) are significantly upregulated or downregulated (see Fig E11 and Table E9 in this article's Online Repository at www.jacionline.org), with 17 of 22 directly interacting with *DOK2* (Fig 3, *B*), thus indicating the biological importance of *DOK2* in AD pathogenesis.

Pathway and gene expression tissue specificity analysis for exome chip variants

To reveal potential differences in terms of biological pathways and involved tissues for exome chip variant (exome chip) and common variants investigated in AD GWAS meta-analysis studies (GWAS) from the EAGLE consortium, we performed gene-set pathway enrichment analysis for curated gene sets and gene ontology terms, as well as tissue-specific gene expression profile analysis,[31] for 53 tissue types and 11,688 samples from the Genotype-Tissue Expression (GTEx) portal, release 7 (see the Methods section in this article's Online Repository). Pathway enrichment analysis identified one significant gene set including

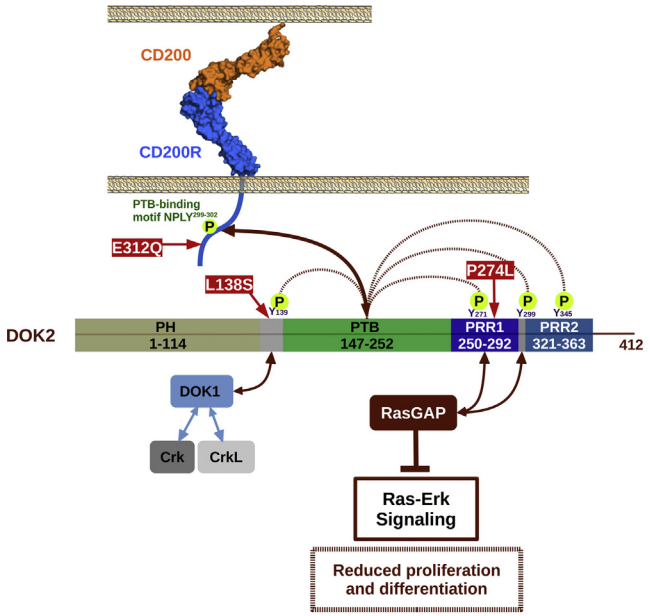


FIG 2. Hypothetical model constructed from protein sequence and structural domain analysis suggesting that missense lead variants rs56094005, rs34215892 (p.L138S and p.P274L; *DOK2*), and rs9865242 (p.E312Q; *CD200R1*) are located near functionally important tyrosine phosphorylation sites and might interfere with CD200/CD200R1 receptor complex and DOK2 function. We assume a simplified illustration of DOK2 function in response to CD200R1 in human myeloid cells[25] in which CD200/CD200R1 binding leads to tyrosine phosphorylation of the NPLY motif in the cytoplasmic tail of CD200R1 and recruitment of the DOK2 adapter protein.[23] DOK2 interaction with CD200R1 leads to DOK2 tyrosine phosphorylation at positions Y271/Y299 (activating RasGAP) and Y139 (activating DOK1), which leads to recruitment of RasGAP and its subsequent activation.[24] Interacting CD200 and CD200R1 extracellular domains are visualized as structural models received from the Protein Databank (ID 4bfi), whereas CD200R1 transmembrane and the cytoplasmic tail are shown as *blue lines*. Variants are highlighted in red to indicate their relative positions within protein domains. The structural effect of protein variants cannot be modelled 3-dimensionally because of the lack of structural templates in the Protein Databank for the concerned regions. Interactions between proteins are visualized as *solid lines* and the following events are visualized as *dashed lines*. Cell membranes anchoring the receptors are illustrated in orange-gray. DOK2 domains were abbreviated as follows: *PH*, Pleckstrin homology domain; *PRR*, proline-rich region; *PTB*, phosphotyrosine-binding domain.

5 genes ($P_{\text{Bonferroni}} < .05$ IL-4 receptor *[IL4R]*, *IL13*, Janus kinase 1 *[JAK1]*, *JAK2*, and tyrosine kinase 2 *[TYK2]*) of the *IL13* pathway for the exome chip in comparison with 8 blood cell–related gene sets (including regulation of immunoglobulin production, B-cell activation, and B cell–mediated immunity) for the EAGLE GWAS data. Recently, AD was characterized as an IL-13–dominant disease based on high-depth RNA-seq transcriptome data of 147 samples from cohorts of patients with AD, patients with psoriasis, and healthy control subjects, with IL-13 being the most distinctive marker for AD.[32] We hypothesize that low-frequency and rare coding variants in genes of the *IL13* pathway are further likely to be associated with AD. In the tissue specificity analysis we tested for relationships between tissue-specific gene expression profiles and variant association statistics from exome chip analyses and GWASs, respectively. For exome chip data, we observed a Bonferroni-corrected significant association ($P < 9.43 \times 10^{-4}$) with 2 skin tissue types (sun-exposed and non–sun-exposed skin; Fig 4, *A*).[31,33] In comparison, the EAGLE GWAS data revealed a significant association with whole blood and spleen tissues (Fig 4, *B*). Thus exome chips variants cumulatively might have a stronger effect on skin tissue gene expression than common variants interrogated in AD consortium GWAS analyses.

Estimation of liability-heritability from exome chip variants

To quantify the overall contribution of exome chip variation to AD risk (ie, estimating the exome-wide contribution of [mainly rare] protein-coding variants to the liability-scale heritability for AD), we used a restricted maximum-likelihood approach to model heritability attributable to genotypic variation across genotyped-only exome chip variants (see the Methods section in this article's Online Repository). We found that common variants from the exome chip (LD pruned, MAF ≥ 5%; see the Methods section in this article's Online Repository) explain 1.04% of heritability on the observed scale (corresponding to 2.93% on the liability scale, assuming a prevalence of 0.14). By decreasing the MAF threshold to 1%, exome chip heritability increased to 1.27% (3.60%). Dividing variants into low-frequency (1% ≤ MAF < 5%) and rare (MAF < 1%), heritability resulted in 0.37% for low-frequency variants (0.94%) and 4.47% for rare variants (12.56%). Furthermore, we estimated the heritability explained by the newly identified lead variants of *DOK2* (rs3215892 and rs56094005) and *CD200R1* (rs9865242). Variants rs3215892, rs56094005, and rs9865242 explain approximately 0.015% (0.041%), 0.004% (0.012%), and 0.013% (0.035%) of the observed scale (liability scale), respectively.

DISCUSSION

We analyzed the association of AD with common, low-frequency, and rare protein-coding variants and implicated 2 novel genes (*DOK2* and *CD200R1*) as contributing to AD risk. Thus far, 14.91% (6.95% excluding *FLG* mutations) of AD liability-scale heritability could be attributed to common lead variants of 31 GWAS loci estimated from a set of 5,775 patients with a clinical diagnosis of AD and assuming a population prevalence of 0.15.[3-9] Recently, Ferreira et al[20] further reported a genome-wide liability-scale heritability estimate of 9.04% from 1.2 million HapMap SNPs (assuming a population prevalence of 0.14) and based on self-reported information from questionnaires. Our newly identified variants (rs3215892, rs56094005, and rs9865242) explain another 0.088% of the variance in liability. Interestingly, approximately 12.56% (SE, 0.74%) of liability-scale heritability is estimated to be explained by rare protein-coding variants (MAF < 1%), thus highlighting the importance of studying rare coding variation in patients with AD. Because these heritability estimates originate from predominantly European study populations, performing GWASs and exome chip studies in more diverse populations is crucial to optimize heritability estimation for AD in the future. Our results encourage future studies along the same path and highlight the importance of studying the effect of protein-coding variants for phenotypically well-defined clinically diagnosed cohorts.

Pathway enrichment analysis of exome chip data revealed that association signals of low-frequency and rare coding variants are enriched in 5 genes of the *IL13* pathway, all of which are targets for novel systemic AD therapeutics in advanced stages of clinical

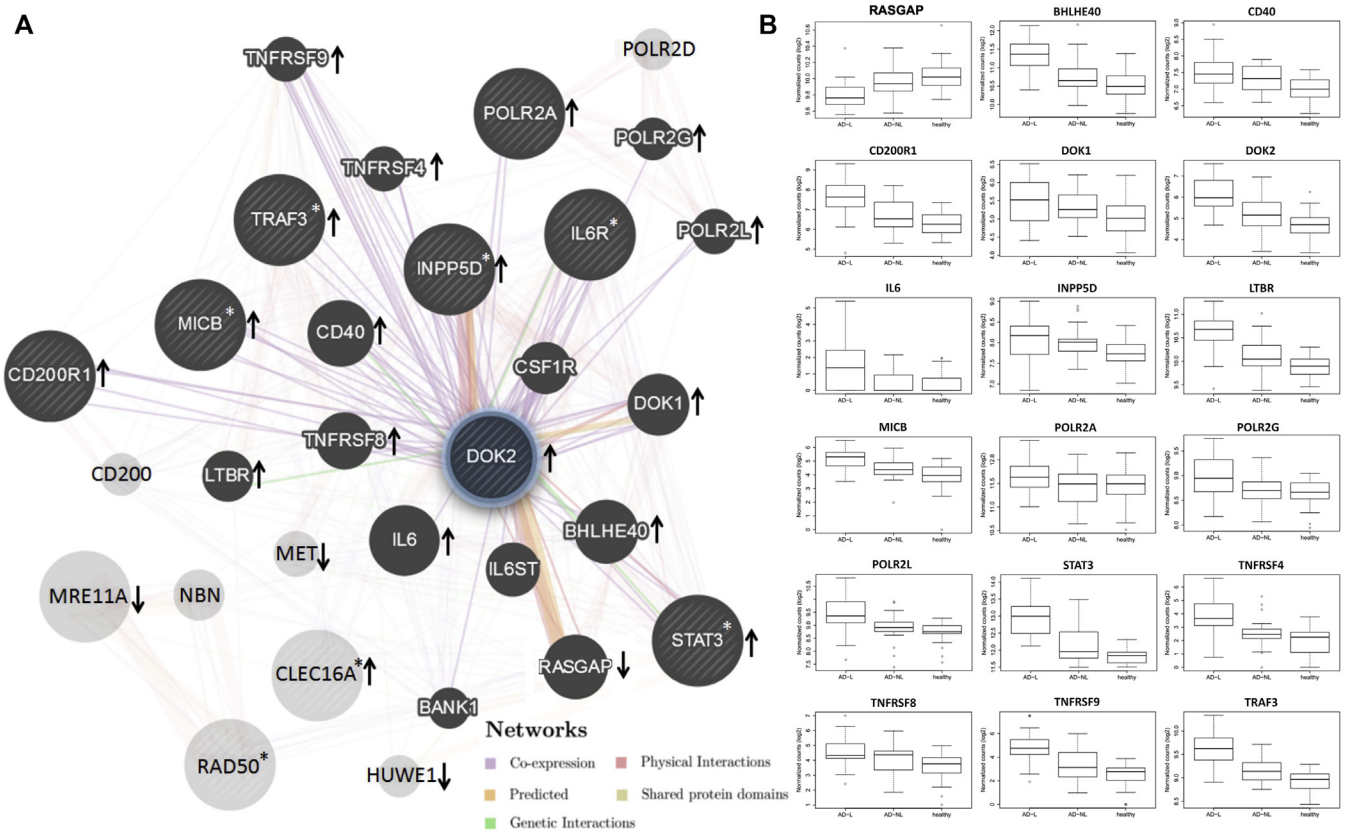


FIG 3. Multiomics-based network interaction analysis unveiled topologically and functionally important AD susceptibility genes and determined *DOK2* to be a central hub node interacting with *CD200R1* and candidate susceptibility genes (marked with *white asterisks*) identified by a previous large AD consortium GWAS. **A,** The majority of core genes (n = 22/30, including *DOK2*) show significantly upregulated and downregulated gene expression levels (depicted by *black upwards and downwards arrows*) in whole-transcriptome mRNA-seq data (AD-L; AD-NL, and healthy, see also Table E9). *Gray-striped nodes* (n = 10, including *DOK2*) represent the most relevant genes from topological network analysis, with 2 *light gray–striped nodes* representing genes not interacting with *DOK2* and 7 *dark gray–striped nodes* representing genes directly interacting with *DOK2*. *Nonstriped nodes* (n = 20) represent genes additionally added by a functional similarity search, with 6 *light gray nonstriped nodes* not directly interacting with *DOK2* and 14 *dark gray nonstriped nodes* directly interacting with *DOK2*. **B,** Seventeen of the significantly upregulated or downregulated genes, depicted with $\log_2$-transformed gene expression counts, interact with *DOK2*.

investigations,[34] supporting the pivotal role of type 2 inflammation in AD pathogenesis.[1] Gene expression tissue profile analysis showed that exome chip variants associated with AD cumulatively have a stronger effect on skin tissue gene expression than common GWAS variants associated with AD, as identified in combination with whole-transcriptome RNA-seq data on lesional, nonlesional, and healthy skin tissue, as well as tissue samples from the GTEx project. In accordance with results from the exome chip study for MS,[13] we observed that the minor allele of low-frequency lead missense variants in *DOK2* is mostly protective (in the context of the OR, Table II).

The newly discovered genes *DOK2* and *CD200R1* have clear immunologic functions, confirming that AD pathogenesis is primarily driven by immune dysregulation. Structural protein domain analysis and topological network and differential gene expression analyses suggest that missense variants in *DOK2* (rs34215892 and rs56094005) in combination with the missense AD risk variant in *CD200R1* (rs9865242) might affect tyrosine phosphorylation sites in *DOK2* and *CD200R1* (Fig 2). *DOK2* belongs to the DOK gene family encoding for 7 different DOK proteins (DOK1-DOK7) involved in signal transduction.[29,35-37] DOK1 and DOK2 are adaptor proteins and mainly expressed in hematopoietic/immune cells[38] and have been implicated to negatively regulate proliferation and constitutive expression of DOK2.[39] Further studies showed that both DOK1 and DOK2 are essential negative regulators of extracellular signal-regulated kinase signaling downstream of Toll-like receptor 4.[40] Mice lacking DOK1 to DOK3 showed significant defects in immune cell development and immune responses.[41] Furthermore, DOK1 and DOK2 play a role in maturation of natural killer cells,[42] numbers of which have been shown to be reduced in patients with AD.

In line with these observations, we observed significantly increased *DOK2* and *CD200R1* mRNA expression in lesional compared with nonlesional skin samples of patients with AD and compared with skin from healthy subjects. Our AD core network (Fig 3) further showed that *DOK2* acts as a central hub gene interacting with *CD200R1*, as well as several candidate AD GWAS susceptibility genes on the cellular level. Our AD core network revealed 21 genes that are directly functionally related to *DOK2*, of which 16 are significantly upregulated

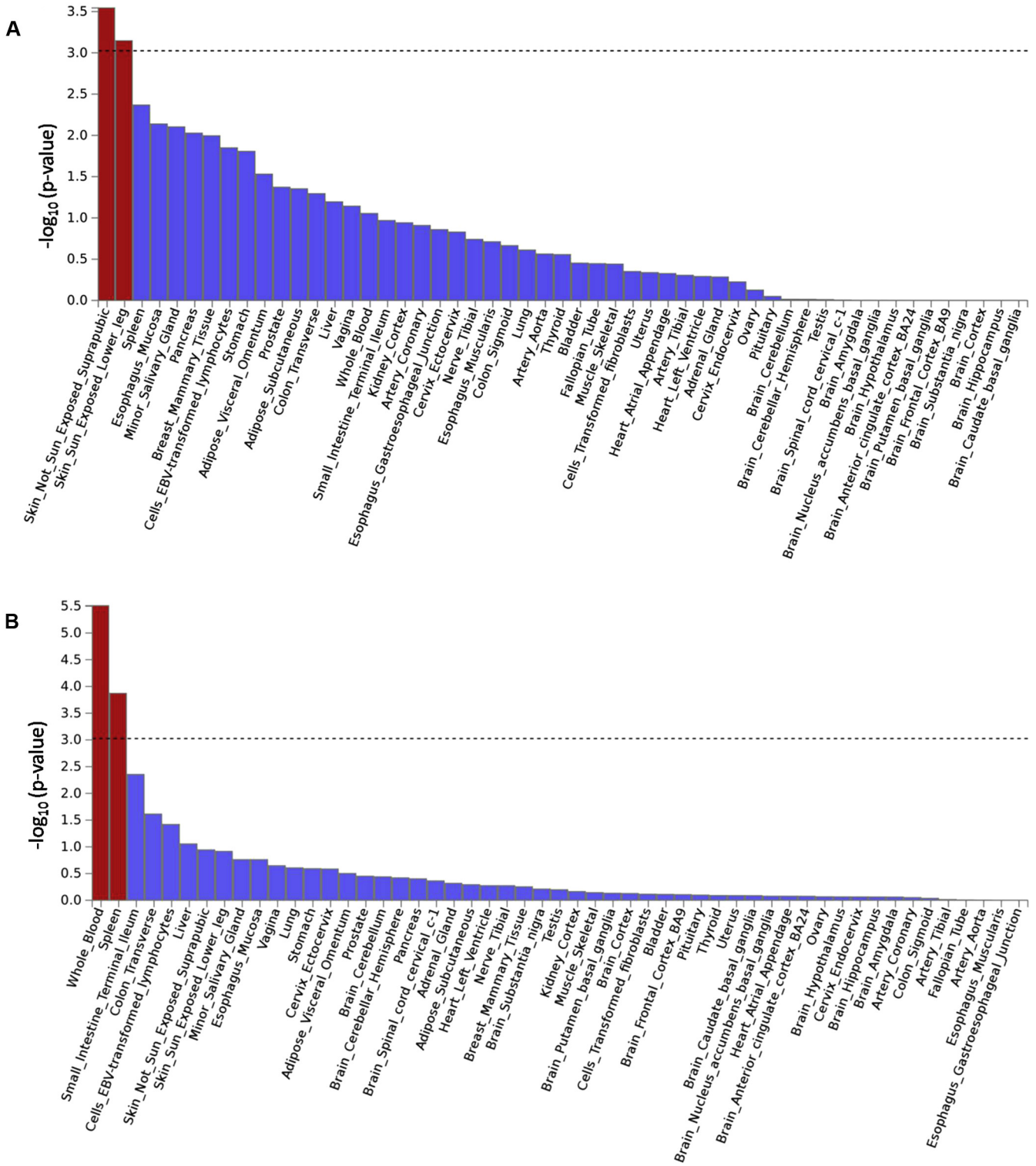


FIG 4. Tissue specificity analysis using 53 tissue types from the GTEx project identifies a significant positive relationship between skin tissue gene expression profiles and exome chip association statistics. Exome chip single variant association score statistics (exome chip data set) and GWAS SNP association summary statistics from the EAGLE eczema consortium (GWAS data set), respectively, were converted to AD-gene association *z* scores by using a multiple linear principal components regression model as implemented in MAGMA,[31] thereby accoutning for LD between SNPs (see the Methods section in this article's Online Repository). We conducted tissue specificity analysis with FUMA[33] to test for a positive relationship between tissue-specific gene expression profiles constructed from 53 tissue types and 11,688 samples from GTEx portal version 7 and AD-gene associations (either from exome chip or GWAS data set) represented by gene-based *z* scores (see the Methods section in this article's Online Repository). We used a Bonferroni-corrected $P < .05/53 = 9.43 \times 10^{-4}$ as the significance cutoff. **A,** The exome chip data set revealed a positive relationship between gene expression profiles from sun-exposed and non–sun-exposed skin tissues and AD-gene associations. **B,** In comparison, the GWAS data set showed an association with whole blood and spleen tissues.

(including *DOK1*) and 1 *(RASGAP)* is significantly downregulated in lesional skin samples of patients with AD. Taking into account the reported inflammatory role of both DOK2 and CD200R1, we conclude that signaling through CD200/CD200R1/DOK2 could be an important new regulatory signaling pathway in AD. Extensive functional studies are required to detect all potential causal variants and thus to specify the contribution of *DOK2* and *CD200R1* to overall disease susceptibility.

We thank Tim Steiert for help with the graphical summary.

Key messages

- **This exome chip analysis for AD identified 2 novel susceptibility genes (*DOK2* and *CD200R1*), and signaling through CD200/CD200R1/DOK2 could be an important new regulatory signaling pathway in patients with AD.**
- **Disease-associated rare coding variants are enriched in 5 genes of the IL-13 pathway, which are targets for novel systemic AD therapeutics in advanced stages of clinical investigations.**
- **In total, 12% of AD heritability is further explained by rare genetic variation, thus highlighting the importance of studying the effect of rare protein-coding variants for AD.**

REFERENCES

1. Weidinger S, Novak N. Atopic dermatitis. Lancet 2016;387:1109-22.
2. Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. Immunol Rev 2011;242:10-30.
3. Schaarschmidt H, Ellinghaus D, Rodriguez E, Kretschmer A, Baurecht H, Lipinski S, et al. A genome-wide association study reveals 2 new susceptibility loci for atopic dermatitis. J Allergy Clin Immunol 2015;136:802-6.
4. Esparza-Gordillo J, Weidinger S, Folster-Holst R, Bauerfeind A, Ruschendorf F, Patone G, et al. A common variant on chromosome 11q13 is associated with atopic dermatitis. Nat Genet 2009;41:596-601.
5. 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;43:690-4.
6. 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;44:1222-6.
7. Paternoster L, Standl M, Chen CM, Ramasamy A, Bonnelykke K, Duijts L, et al. Meta-analysis of genome-wide association studies identifies three new risk loci for atopic dermatitis. Nat Genet 2011;44:187-92.
8. Ellinghaus D, Baurecht H, Esparza-Gordillo J, Rodriguez E, Matanovic A, Marenholz I, et al. High-density genotyping study identifies four new susceptibility loci for atopic dermatitis. Nat Genet 2013;45:808-12.
9. Paternoster L, Standl M, Waage J, Baurecht H, Hotze M, Strachan DP, et al. Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. Nat Genet 2015;47:1449-56.
10. Irvine AD, McLean WH, Leung DY. Filaggrin mutations associated with skin and allergic diseases. N Engl J Med 2011;365:1315-27.
11. Tsunemi Y, Saeki H, Nakamura K, Sekiya T, Hirai K, Kakinuma T, et al. Interleukin-13 gene polymorphism G4257A is associated with atopic dermatitis in Japanese patients. J Dermatol Sci 2002;30:100-7.
12. Esparza-Gordillo J, Schaarschmidt H, Liang L, Cookson W, Bauerfeind A, Lee-Kirsch MA, et al. A functional IL-6 receptor (IL6R) variant is a risk factor for persistent atopic dermatitis. J Allergy Clin Immunol 2013;132:371-7.
13. International Multiple Sclerosis Genetics Consortium, Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, et al. Low-frequency and rare-coding variation contributes to multiple sclerosis risk. Cell 2018;175:1679-1687, e7.
14. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285-91.
15. Nothlings U, Krawczak M. [PopGen. A population-based biobank with prospective follow-up of a control group]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2012;55:831-5.
16. Wichmann HE, Gieger C, Illig T, Group MKS. KORA-gen–resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen 2005;67(suppl 1):S26-30.
17. Schmermund A, Mohlenkamp S, Stang A, Gronemeyer D, Seibel R, Hirche H, et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. Risk factors, evaluation of coronary calcium and lifestyle. Am Heart J 2002;144:212-8.
18. Volzke H, Alte D, Schmidt CO, Radke D, Lorbeer R, Friedrich N, et al. Cohort profile: the study of health in Pomerania. Int J Epidemiol 2011;40:294-307.
19. Bisgaard H. The Copenhagen Prospective Study on Asthma in Childhood (COPSAC): design, rationale, and baseline data from a longitudinal birth cohort study. Ann Allergy Asthma Immunol 2004;93:381-9.
20. Ferreira MA, Vonk JM, Baurecht H, Marenholz I, Tian C, Hoffman JD, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. Nat Genet 2017;49:1752-7.
21. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J, et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 2013;45:1238-43.
22. Fehrmann RS, Jansen RC, Veldink JH, Westra HJ, Arends D, Bonder MJ, et al. Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. PLoS Genet 2011;7:e1002197.
23. Mihrshahi R, Barclay AN, Brown MH. Essential roles for Dok2 and RasGAP in CD200 receptor-mediated regulation of human myeloid cells. J Immunol 2009;183:4879-86.
24. Mihrshahi R, Brown MH. Downstream of tyrosine kinase 1 and 2 play opposing roles in CD200 receptor signaling. J Immunol 2010;185:7216-22.
25. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, et al. The NIH Roadmap Epigenomics Mapping Consortium. Nat Biotechnol 2010;28:1045-8.
26. Mashima R, Hishida Y, Tezuka T, Yamanashi Y. The roles of Dok family adapters in immunoreceptor signaling. Immunol Rev 2009;232:273-85.
27. Holmannova D, Kolackova M, Kondelkova K, Kunes P, Krejsek J, Andrys C. CD200/CD200R paired potent inhibitory molecules regulating immune and inflammatory responses; Part I: CD200/CD200R structure, activation, and function. Acta Medica (Hradec Kralove) 2012;55:12-7.
28. Zhang S, Cherwinski H, Sedgwick JD, Phillips JH. Molecular mechanisms of CD200 inhibition of mast cell activation. J Immunol 2004;173:6786-93.
29. Boulay I, Nemorin JG, Duplay P. Phosphotyrosine binding-mediated oligomerization of downstream of tyrosine kinase (Dok)-1 and Dok-2 is involved in CD2-induced Dok phosphorylation. J Immunol 2005;175:4483-9.
30. Schrambach S, Ardizzone M, Leymarie V, Sibilia J, Bahram S. In vivo expression pattern of MICA and MICB and its relevance to auto-immunity and cancer. PLoS One 2007;2:e518.
31. de Leeuw CA, Mooij JM, Heskes T, Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol 2015;11:e1004219.
32. Tsoi LC, Rodriguez E, Degenhardt F, Baurecht H, Wehkamp U, Volks N, et al. Atopic dermatitis is an IL-13 dominant disease with greater molecular heterogeneity compared to psoriasis. J Invest Dermatol 2019;139:1480-9.
33. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun 2017;8:1826.
34. Weidinger S, Beck LA, Bieber T, Kabashima K, Irvine AD. Atopic dermatitis. Nat Rev Dis Primers 2018;4:1.
35. Abramson J, Rozenblum G, Pecht I. Dok protein family members are involved in signaling mediated by the type 1 Fcepsilon receptor. Eur J Immunol 2003;33:85-91.
36. Gerard A, Favre C, Garcon F, Nemorin JG, Duplay P, Pastor S, et al. Functional interaction of RasGAP-binding proteins Dok-1 and Dok-2 with the Tec protein tyrosine kinase. Oncogene 2004;23:1594-8.
37. Yasuda T, Bundo K, Hino A, Honda K, Inoue A, Shirakata M, et al. Dok-1 and Dok-2 are negative regulators of T cell receptor signaling. Int Immunol 2007;19:487-95.
38. Di Cristofano A, Carpino N, Dunant N, Friedland G, Kobayashi R, Strife A, et al. Molecular cloning and characterization of p56dok-2 defines a new family of RasGAP-binding proteins. J Biol Chem 1998;273:4827-30.

39. Gugasyan R, Quilici C,IST, Grail D, Verhagen AM, Roberts A, et al. Dok-related protein negatively regulates T cell development via its RasGTPase-activating protein and Nck docking sites. J Cell Biol 2002;158:115-25.
40. Shinohara H, Inoue A, Toyama-Sorimachi N, Nagai Y, Yasuda T, Suzuki H, et al. Dok-1 and Dok-2 are negative regulators of lipopolysaccharide-induced signaling. J Exp Med 2005;201:333-9.
41. Mashima R, Arimura S, Kajikawa S, Oda H, Nakae S, Yamanashi Y. Dok adaptors play anti-inflammatory roles in pulmonary homeostasis. Genes Cells 2013;18:56-65.
42. Celis-Gutierrez J, Boyron M, Walzer T, Pandolfi PP, Jonjic S, Olive D, et al. Dok1 and Dok2 proteins regulate natural killer cell development and function. EMBO J 2014;33:1928-40.