





Aotearoa New Zealand Māori and Pacific Population-amplified Gout Risk Variants: *CLNK* Is a Separate Risk Gene at the *SLC2A9* Locus

Aichang Ji¹ , Amara Shaukat², Riku Takei³ , Matthew Bixley², Murray Cadzow² , Ruth K. Topless², Tanya J. Major², Amanda Phipps-Green², Marilyn E. Merriman², Jennie Harré Hindmarsh⁴, Lisa K. Stamp⁵, Nicola Dalbeth⁶, Changgui Li¹, and Tony R. Merriman⁷ 

ABSTRACT. *Objective.* The Māori and Pacific (Polynesian) population of Aotearoa New Zealand has a high prevalence of gout. Our aim was to identify potentially functional missense genetic variants in candidate inflammatory genes amplified in frequency that may underlie the increased prevalence of gout in Polynesian populations. *Methods.* A list of 712 inflammatory disease-related genes was generated. An in silico targeted exome set was extracted from whole genome sequencing data in people with gout of various ancestral groups (Polynesian, European, East Asian; n = 55, 780, 135, respectively) to identify Polynesian-amplified common missense variants (minor allele frequency > 0.05). Candidate functional variants were tested for association with gout by multivariable-adjusted regression analysis in 2528 individuals of Polynesian ancestry. *Results.* We identified 26 variants common in the Polynesian population and uncommon in the European and East Asian populations. Three of the 26 population-amplified variants were nominally associated with the risk of gout (*rs1635712* [*KLAA0319*], $OR_{meta} = 1.28$, $P_{meta} = 0.03$; *rs16869924* [*CLNK*], $OR_{meta} = 1.37$, $P_{meta} = 0.002$; *rs2070025* [fibrinogen A alpha chain (*FGA*)], $OR_{meta} = 1.34$, $P_{meta} = 0.02$). The *CLNK* variant, within the established *SLC2A9* gout locus, was genetically independent of the association signal at *SLC2A9*. *Conclusion.* We provide nominal evidence for the existence of population-amplified genetic variants conferring risk of gout in Polynesian populations. Polymorphisms in *CLNK* have previously been associated with gout in other populations, supporting our evidence for the association of this gene with gout.

Key Indexing Terms: association, gene, gout, Māori, Pacific, Polynesian

This research was supported by the Health Research Council of New Zealand (Grant 14/527).

¹A. Ji, PhD, Research Fellow, C. Li, PhD, Professor, Shandong Provincial Key Laboratory of Metabolic Diseases, the Affiliated Hospital of Qingdao University, Qingdao, China; ²A. Shaukat, MSc, Doctoral Student, M. Bixley, MSc, Assistant Research Fellow, M. Cadzow, PhD, Research Fellow, R.K. Topless, BSc, Assistant Research Fellow, T.J. Major, PhD, Research Fellow, A. Phipps-Green, MSc, Assistant Research Fellow, M.E. Merriman, BSc, Research Assistant, Department of Biochemistry, University of Otago, Dunedin, New Zealand; ³R. Takei, MSc, Scientist, Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Birmingham, USA; ⁴J. Harré Hindmarsh, PhD, Research Coordinator, Ngāti Porou Hauora Charitable Trust, Te Puia Springs, Tairāwhiti East Coast, New Zealand; ⁵L.K. Stamp, PhD, Professor, Department of Medicine, University of Otago, Christchurch, New Zealand; ⁶N. Dalbeth, MD, Professor, Department of Medicine, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand; ⁷T.R. Merriman, BSc, Research Assistant, Shandong Provincial Key Laboratory of Metabolic Diseases, the Affiliated Hospital of Qingdao University, Qingdao, China, Department of Biochemistry, University of Otago, Dunedin, New Zealand, and Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, Birmingham, USA.

A. Ji and A. Shaukat contributed equally to this work.

The authors declare no conflict of interest relevant to this article.

Address correspondence to T.R. Merriman, School of Biomedical Sciences, Department of Biochemistry, 710 Cumberland Street, Dunedin, Otago 9054, New Zealand. Email: tony.merriman@otago.ac.nz.

Accepted for publication June 11, 2021.

Gout is an inflammatory arthritis triggered by an immune response to monosodium urate (MSU) crystals inside the joints when circulating urate levels are elevated (hyperuricemia). In New Zealand in 2016, the prevalence of gout was 2- to 3.5-fold higher in Māori (8%) and Pacific (14%) people than those of non-Māori and non-Pacific ancestry (4%).¹ While structural inequities contribute to the increased prevalence,² research on population-specific genetic variants associated with gout is important to provide insights into the disparities between different populations.

Candidate gene studies and genome-wide association studies (GWAS) have identified hundreds of variants associated with urate levels and gout.^{3,4,5,6,7,8} Many of the identified loci include genes encoding renal and gut urate transporters, such as *SLC2A9*, *ABCG2*, *SLC17A1*, *SLC22A11*, and *SLC22A12*, that affect the risk of gout through regulation of excretion of urate.⁹ However, although hyperuricemia is essential in the pathogenesis of gout, it is not sufficient for the development of clinical gout.¹⁰ Activation of the NLRP3 inflammasome is crucial for the production of interleukin (IL)-1 β in response to MSU crystals and the triggering of the inflammatory cascade. There is a need to understand the genetic factors influencing the development of gout in the presence of hyperuricemia; however, research progress has been limited. Several genes potentially involved in

the progression from hyperuricemia to gout have been identified, including variants in *PPARGC1B*, *CARD8*, *IL-1 β* , *CD14*, *TLR4*, *APOA1*, *APOC3*, and *P2RX7*.^{11,12,13,14,15,16}

There is increasing evidence for the existence of population-specific variants for gout. For example, GWAS have found 3 East Asian-specific signals on chromosome 11 (*SLC22A9*, *PLA2G16*, *AIP*),³ ABCG2 Q126X in the Japanese population,¹⁷ IL-37 N182S and ABCC4 P1036L variants in Polynesian populations,^{18,19} and a Polynesian-specific mitochondrial haplogroup associated with gout.²⁰ Combined with the existence of Polynesian-specific variants in the *CREBRF* and *CETP* genes that associate with other metabolic phenotypes,^{21,22} the identification of further Polynesian-specific variants in candidate gout genes may yield new insights into the pathogenesis of gout in Polynesian populations.

In this study, we focused on genes implicated in inflammatory diseases and identified 26 missense variants common in Polynesian populations and uncommon in the European and East Asian populations from whole genome sequencing (WGS) data. These variants were tested for association with gout in the Māori and Pacific populations of Aotearoa New Zealand.

METHODS

Study cohorts. Demographic and biochemical characteristics of the participants are summarized in Table 1. Gout cases were diagnosed according to the 1977 American Rheumatology Association preliminary criteria.²³ WGS data of people with gout (Polynesian, European, East Asian; n = 55, 780, 135, respectively) from clinical trials for urate-lowering therapy^{24,25} and a gout and related diseases study²¹ were included. Follow-up association analysis was done in a cohort genome-wide genotyped on the Human CoreExome v24 (Illumina) array.²¹ For the East Polynesian (EP; New Zealand Māori, Cook Island Māori) group, 800 people with gout and 810 without gout participated, and for the West Polynesian (WP; Samoa, Tonga, Tuvalu, Tokelau, Niue) group, 481 people with gout and 437 without gout participated. A separate Māori sample set from the *rohe* (area) of the Ngāti Porou *iwi* (tribe) of the Tairāwhiti (East Coast, the North Island of New Zealand) region was also included in the EP group. This sample set was recruited in collaboration with the Ngāti Porou Hauora (Health Service) Charitable Trust. Because of genetic differences between the Eastern and Western Polynesian populations,^{21,26} the EP and WP groups were analyzed separately.

All participants provided written informed consent and ethical approval was obtained from the New Zealand Multiregional Ethics Committee (MEC05/10/130), the Northern Y Region Health Research Ethics Committee (Ngāti Porou Hauora Charitable Trust study; NTY07/07/074), and, in the US, the Schulman Central Institutional Review Board (201102890).

Identification of candidate genes. GWAS for immune-mediated conditions with an autoinflammatory aspect such as inflammatory bowel disease (Crohn disease and ulcerative colitis), systemic lupus erythematosus (SLE), systemic sclerosis, rheumatoid arthritis (RA), ankylosing spondylitis, and gout in European and East Asian populations were sourced from the literature prior to April 30, 2019 (Supplementary Material, available from the authors on request). From the various GWAS, genetically associated loci were defined using either the summary statistics, or reported lead single-nucleotide polymorphisms (SNPs). If summary statistics were available, the less significant variants ($P > 0.0001$) were filtered out and associated regions on each chromosome were clustered using a density-based clustering algorithm in R from the package *dbSCAN*.²⁷ Each region was scored and ranked based on the minimum P value within the region, and the top 20 regions per GWAS were selected. Candidate genes selected were in the ± 100 kb region from the lead SNP. Four hundred fifty-one genes were identified from the GWAS approach.

Three hundred ninety-two genes were identified from a literature search using keywords “inflammation genes” or “genes involved inflammatory pathways” in PubMed, Google Scholar, and ScienceDirect. Genes linked to monogenic disorders with an innate immune basis were also included. There were 712 genes remaining after removing duplicated genes within the combined GWAS and literature search list (Supplementary Material, available from the authors on request).

WGS and in silico extraction. Genomes were sequenced at the Garvan Institute of Medical Research, Kinghorn Cancer Centre, New South Wales, Australia. The platform used was 30x WGS (TruSeq Nano) v2.5. The FASTQ format files of sequences were aligned to the human reference genome (GRCh37) and variant called following an implementation of the Genome Analysis ToolKit (GATK) best practices²⁸ using Burrows-Wheeler,²⁹ Picardtools (<https://broadinstitute.github.io/picard/>), and GATK v3.6.0.³⁰

Only missense, nonsense, splicing variants and insertion, and/or deletions (indels) within the candidate genes that passed standard quality control criteria (mapping quality score > 20 ; heterozygosity > 0.001 ; indel heterozygosity = 1.25×10^{-4} ; genotype quality > 10 ; mapping quality > 30 ; and those that had high coverage of reads and genotype quality [> 90]) were included for further analysis. Genetic variations found in the discovery phase were prioritized based on those that were common (minor allele frequency [MAF] > 0.05) in the Polynesian sample set but uncommon (MAF < 0.01) in the European and East Asian sample sets. The ANNOVAR tool was used to annotate the genes and protein structural consequence of the variants.³¹ Only missense variants also present on the Human CoreExome v24 (Illumina) array were included in the association analysis.

Statistical analysis. Multivariable-adjusted logistic regression analysis and linear regression analysis were used to test for association with gout and serum urate concentration (in controls), respectively, using the statistical software PLINK version 1.9.³² Allelic ORs and β (mmol/L) were both calculated for each SNP by adjusting for age, sex, and principal components (derived from the genome-wide genotype data as previously described)²¹ in the regression models. Individuals with missing data for any variable were

Table 1. Demographic details of the datasets.

	Eastern Polynesian		Western Polynesian	
	Gout	Controls	Gout	Controls
Total participants	800	810	481	437
Male, n (%)	603 (75.4)	315 (38.9)	421 (87.5)	226 (51.7)
Age, yrs	56.4 \pm 12.6 (21–100)	45.5 \pm 15.6 (15–94)	47.3 \pm 12.7 (17–80)	33.7 \pm 15.9 (14–80)
BMI, kg/m ²	36.3 \pm 8.4 (18.2–83.7)	32.6 \pm 8.5 (16.2–80.0)	36.6 \pm 7.6 (21.2–79.6)	32.5 \pm 7.4 (19.1–71.1)
Serum urate, mmol/L	0.40 \pm 0.13 (0.05–0.86)	0.35 \pm 0.10 (0.04–0.78)	0.44 \pm 0.14 (0.08–0.85)	0.35 \pm 0.10 (0.05–0.73)

Data are shown as mean \pm SD (range) unless otherwise stated.

excluded. A P value ≤ 0.05 was set for nominal statistical significance. The dataset was not sufficiently powered to correct for multiple testing ($n = 26$ variants, separate EP and WP analyses) and, in the absence of an available dataset for replication, the association evidence presented here can only be considered as nominal with external information required for validation (as we were able to do for the *CLNK* locus). Inverse variance-weighted fixed-effect metaanalysis was performed using the software METAL³³ to combine independent sample sets. Heterogeneity between combined sample sets was determined through the Q statistic, and for variants with significant evidence for heterogeneity ($P_{\text{Het}} < 0.05$), the fixed-effect model was replaced with a random-effect model. A LocusZoom plot including calculations of linkage disequilibrium (LD, measured in r^2)³⁴ was produced using the entire Polynesian sample set. The Combined

Annotation-Dependent Depletion (CADD) score of the significant SNPs was sourced.³⁵ CADD scores represent rankings compared to other protein-coding variants within the entire genome. For example, a CADD score of ≥ 10 indicates 10% of variants in the genome are predicted to be more deleterious, whereas a score of ≥ 20 indicates 1% are predicted to be more deleterious.

RESULTS

The study design is summarized in Figure 1. A total of 26 missense variants were extracted from the Polynesian, European, and East Asian WGS data within the exome of the 712 candidate genes and filtered according to being common (MAF > 0.05) in

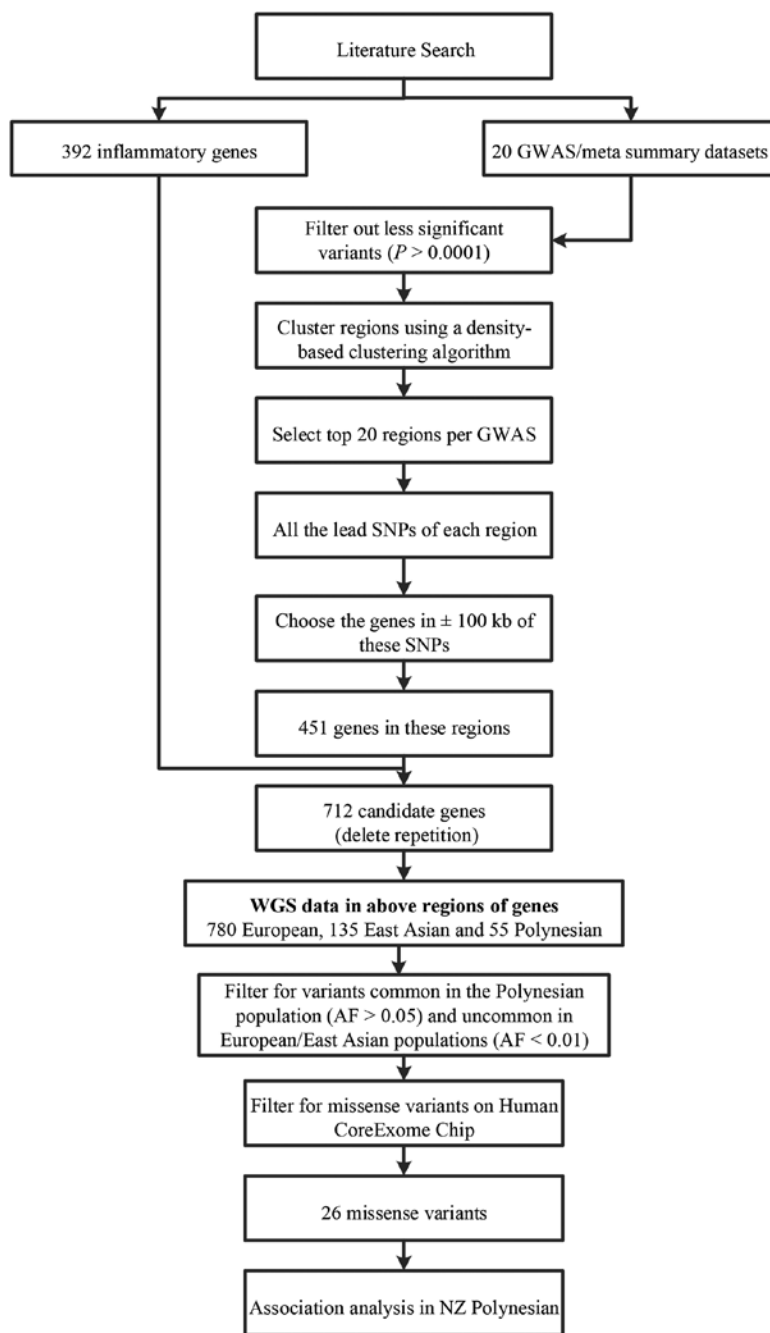


Figure 1. Analysis workflow. AF: allele frequency; GWAS: genome-wide association study; NZ: New Zealand; SNP: single-nucleotide polymorphism; WGS: whole genome sequencing.

the Polynesian population but uncommon (MAF < 0.01) in the other populations (Table 2).

Metaanalysis of the Polynesian EP and WP subgroups supported association of *rs1635712* (*KIAA0319*), *rs16869924* (*CLNK*, cytokine-dependent hematopoietic cell linker), and *rs2070025* (fibrinogen A α chain [*FGA*]) with the risk of gout (OR 1.28 [$P = 0.03$], OR 1.37 [$P = 0.002$], and OR 1.34 [$P = 0.02$], respectively; Table 3, Figure 2). *KIAA0319* and *CLNK* were included in our study as GWAS loci, associated with SLE and RA, respectively (Supplementary Material, available from the authors on request). *FGA* was included in our study from the literature search as a candidate gene

for inflammation (Supplementary Material). The variant in *KIAA0319* (*rs1635712*) had a derived T allele frequency of 11.4% in EP and 14.4% in WP controls (Table 3). For *rs16869924* (*CLNK*) the derived C allele frequency was 25.0% in EP and 11.6% in WP controls (Table 3). For *rs2070025* (*FGA*) the frequency of the derived C allele was 13.0% in EP and 4.1% in WP controls (Table 3). The CADD score was 4.8 for *rs16869924*, 7.4 for *rs1635712*, and 14.6 for *rs2070025*.

Rs16869924 (*CLNK*) is located 526 kb downstream from *SLC2A9*, a locus with strong association with gout in Polynesian populations.⁷ Therefore, we tested for the possibility of linkage disequilibrium between *rs16869924* and gout-associated

Table 2. The 26 missense variants selected to test for association with gout.

SNP	Gene	Ref/Alt Allele	Chr	Position (GRCh37)	Amino Acid Change	MAF		MAF		
						gnomADv2 Exomes	gnomADv2 Genomes	Polynesian n = 55	European n = 780	East Asian n = 135
rs1635712	<i>KIAA0319L</i>	C/T	1	35944752	p.Gly243Asp	0.0056	0.0123	0.1182	0.0006	6.66e-05
rs148776402	<i>CLC42</i>	T/A	1	86905915	p.Leu430Ile	0.0002	0.0002	0.1545	0	0
rs200123911	<i>HIPK1</i>	C/T	1	114510534	p.Ser449Leu; p.Ser469Leu; p.Ser843Leu	9.156e-05	0.0001	0.1091	0	0.0001
rs148561683	<i>PTPRC</i>	A/C	1	198672489	p.Thr55Pro; p.Thr216Pro	0.0005	0.0008	0.0636	0	0
rs150493510	<i>LY75</i>	T/G	2	160755340	p.Lys109Gln	0.0002	0.0001	0.0818	0	0
rs16869924	<i>CLNK</i>	T/C	4	10567732	p.Ser65Gly	0.0054	0.0069	0.1727	0.0012	6.66e-05*
rs2070025	<i>FGA</i>	T/C	4	155511824	p.Ile6Val	0.0074	0.0051	0.2000	0	0*
rs146544425	<i>PRKAA1</i>	T/C	5	40764876	p.Lys429Arg; p.Lys444Arg	0.0005	0.0004	0.1091	0.0006	0*
rs148648401	<i>TTC37</i>	G/C	5	94860194	p.Thr476Arg	0.0031	0.0027	0.2364	0	0*
rs118132132	<i>ERAP1</i>	T/G	5	96130769	p.Ile299Leu	0.0011	0.0005	0.1909	0.0006	0
rs2071471	<i>TAP2</i>	G/A	6	32784645	p.Leu673Phe	0.0087	0.0059	0.1273	0.0006	0.0005*
rs148570036	<i>TAGAP</i>	T/C	6	159457034	p.His611Arg; p.His496Arg; p.His674Arg	0.0027	0.0024	0.1909	0	0*
rs199943004	<i>SNAPC4</i>	C/T	9	139277766	p.Ala619Thr	0.0005	0.0004	0.1321	0	0
rs201030657	<i>SNAPC4</i>	C/T	9	139277793	p.Gly610Ser	0.0006	0.0004	0.1321	0	0
rs201808826	<i>SDCCAG3</i>	G/A	9	139299121	p.His271Tyr; p.His321Tyr; p.His344Tyr	0.0005	0.0004	0.1636	0	0
rs11555379	<i>PRR5L</i>	A/G	11	36440836	p.Ile12Val; p.Ile93Val	0.0094	0.0144	0.2545	0.0012	0.0016*
rs114820168	<i>IRAK4</i>	C/T	12	44177510	p.Arg267Cys; p.Arg188Cys; p.Arg391Cys	0.0002	0.0003	0.0909	0	0
rs111988600	<i>SLC26A10</i>	G/C	12	58016690	p.Lys304Asn	0.0018	0.0014	0.1636	0.0032	0.0013
rs112603554	<i>SLC26A10; B4GALNT1</i>	T/G	12	58016890	p.Asn341Lys	0.0035	0.0026	0.1636	0.0050	0.0028
rs147301375	<i>B4GALNT1</i>	C/T	12	58022618	p.Ala239Thr; p.Ala294Thr	0.0018	0.0015	0.1636	0.0031	0.0013
rs8176344	<i>CYP27B1</i>	C/G	12	58159173	p.Val166Leu	0.0099	0.0019	0.12963	0.0038	0.0017
rs201730712	<i>APOBR</i>	A/G	16	28508491	p.Asp710Gly	0.0024	0.0006	0.02727	0.0012	0.0005
rs74530614	<i>FEAR1</i>	C/A	19	35843347	p.Ser298Tyr	0.0040	0.0018	0.2037	0.0019	0.0002*
rs78635115	<i>ALDH16A1</i>	T/C	19	49964111	p.Ser178Pro	0.0004	0.0005	0.2364	0.0006	0
rs77072552	<i>NLRP7</i>	C/T	19	55449446	p.Val69Ile; p.Val671Ile	0.0035	0.0026	0.1273	0.0019	0.0002*
rs104895509	<i>NLRP7</i>	C/T	19	55452829	p.Cys84Tyr; p.Cys84Tyr	0.0004	0.0003	0.1364	0	6.68e-05

* MAF > 0.01 in East Asian 1000 genomes. Alt: alternative; Chr: chromosome; MAF: minor allele frequency; Ref: reference; SNP: single-nucleotide polymorphism.

Table 3. Association analysis of the 26 missense variants with gout.

SNP	Gene	Eastern Polynesian			Western Polynesian			P_{meta}	OR (95% CI)	
		MAF, Cases	MAF, Controls	P	MAF, Cases	MAF, Controls	P			
rs1635712	<i>KIAA0319L</i>	0.136	0.114	0.07	0.164	0.144	0.22	1.23 (0.89–1.70)	0.03	1.28 (1.03–1.59)
rs148776402	<i>CLCA2</i>	0.148	0.162	0.08	0.045	0.029	0.76	1.11 (0.58–2.09)	0.14	0.83 (0.65–1.06)
rs200123911	<i>HIPK1</i>	0.097	0.102	0.60	0.040	0.039	0.25	0.70 (0.38–1.29)	0.44	0.86 (0.65–1.15)
rs148561683	<i>PTPRC</i>	0.060	0.059	0.82	0.127	0.100	0.33	1.20 (0.83–1.75)	0.39	1.13 (0.86–1.49)
rs150493510	<i>LY75</i>	0.031	0.032	0.83	0.188	0.155	0.16	1.26 (0.92–1.73)	0.18	1.21 (0.92–1.61)
rs16869924	<i>CLNK</i>	0.261	0.250	0.08	0.177	0.116	0.002	1.75 (1.22–2.50)	0.002	1.37 (1.13–1.68)
rs2070025	<i>FGA</i>	0.176	0.130	0.06	0.079	0.041	0.12	1.52 (0.89–2.61)	0.02	1.34 (1.05–1.71)
rs146544425	<i>PRKAA1</i>	0.121	0.114	0.57	0.240	0.217	0.49	1.11 (0.83–1.48)	0.37	1.10 (0.89–1.36)
rs148648401	<i>TTC37</i>	0.201	0.194	0.72	0.292	0.252	0.13	1.24 (0.94–1.62)	0.45	1.07 (0.89–1.29)
rs118132132	<i>ERAP1</i>	0.264	0.246	0.59	0.172	0.143	0.31	1.19 (0.85–1.66)	0.31	1.10 (0.91–1.33)
rs2071471	<i>TAP2HLA</i>	0.085	0.080	0.49	0.116	0.132	0.55	0.89 (0.62–1.29)	0.36	0.88 (0.67–1.16)
rs148570036	<i>TAGAP</i>	0.283	0.284	0.52	0.114	0.090	0.64	0.91 (0.61–1.36)	0.43	0.92 (0.76–1.13)
rs199943004 ^a	<i>SNAPC4</i>	0.139	0.131	0.56	0.140	0.154	0.07	0.73 (0.52–1.02)	0.30	0.84 (0.67–1.04)
rs201030657 ^a	<i>SNAPC4</i>	0.140	0.131	0.57	0.140	0.155	0.06	0.72 (0.52–1.01)	0.28	0.84 (0.67–1.04)
rs201808826 ^a	<i>SDCAG3</i>	0.140	0.125	0.80	0.150	0.158	0.15	0.78 (0.56–1.09)	0.26	0.88 (0.71–1.10)
rs11555379	<i>PRR5L</i>	0.267	0.239	0.49	0.153	0.158	0.24	0.82 (0.58–1.15)	0.93	0.99 (0.82–1.20)
rs114820168	<i>IRAK4</i>	0.052	0.052	0.45	0.143	0.143	> 0.99	1.00 (0.70–1.42)	0.64	1.07 (0.81–1.41)
rs111988600 ^a	<i>SLC26A10</i>	0.173	0.168	0.41	0.174	0.174	0.71	1.06 (0.77–1.46)	0.39	1.09 (0.89–1.34)
rs112603554 ^a	<i>SLC26A10; B4GALNT1</i>	0.174	0.170	0.42	0.174	0.174	0.71	1.06 (0.77–1.46)	0.39	1.09 (0.89–1.34)
rs147301375 ^a	<i>B4GALNT1</i>	0.173	0.168	0.41	0.174	0.175	0.74	1.06 (0.77–1.45)	0.40	1.09 (0.89–1.34)
rs8176344	<i>CYP27B1</i>	0.149	0.147	0.63	0.142	0.132	0.60	1.10 (0.78–1.54)	0.48	1.08 (0.87–1.34)
rs201730712	<i>APOBR</i>	0.003	0.004	0.73	0.135	0.112	0.36	1.18 (0.83–1.67)	0.61	1.16 (0.82–1.63)
rs74530614	<i>FEAR1</i>	0.230	0.227	0.87	0.182	0.156	0.50	1.12 (0.81–1.54)	0.79	1.03 (0.85–1.24)
rs78635115	<i>ALDH16A1</i>	0.234	0.224	0.99	0.240	0.219	0.86	1.03 (0.76–1.38)	0.93	1.01 (0.84–1.21)
rs77072552	<i>NLRP7</i>	0.241	0.212	0.18	0.222	0.176	0.34	1.16 (0.86–1.56)	0.10	1.17 (0.97–1.41)
rs104895509	<i>NLRP7</i>	0.086	0.080	0.78	0.096	0.083	0.78	1.06 (0.70–1.61)	0.97	0.99 (0.76–1.30)

^a Variants at the *SNAPC4* locus were in strong linkage disequilibrium (*rs199943004*, *rs201030657*, $r^2 = 1$; *rs199943004*, *rs201808826*, $r^2 = 0.95$). Variants at the *SLC26A10* locus were in strong linkage disequilibrium (*rs111988600*, *rs112603554*, $r^2 = 0.99$; *rs111988600*, *rs147301375*, $r^2 = 0.99$). MAF: minor allele frequency; SNP: single-nucleotide polymorphism.

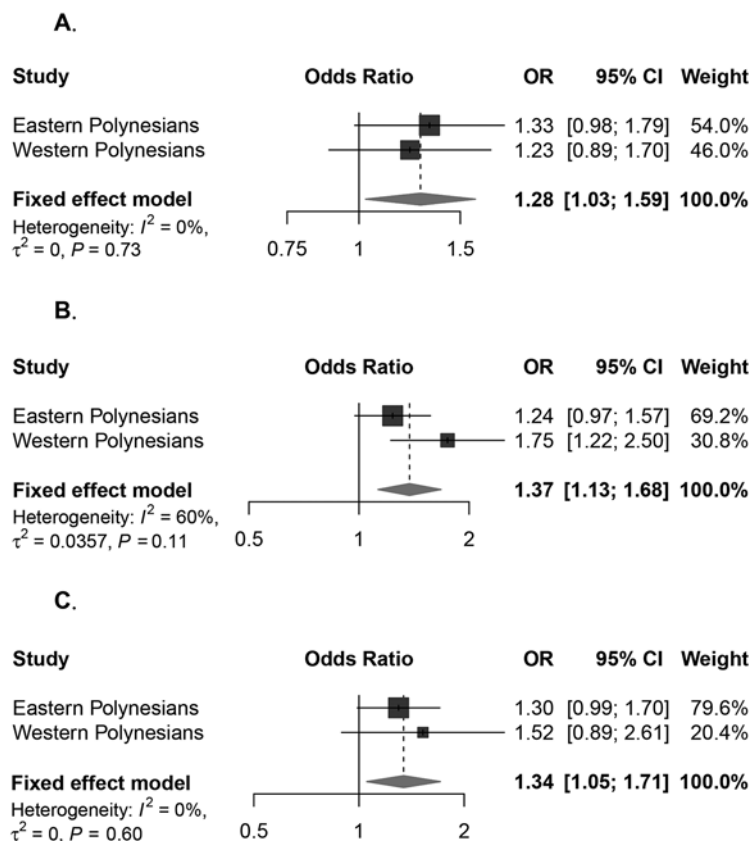


Figure 2. Forest plot of metaanalysis for (A) *rs1635712* (*KIAA0319*), (B) *rs16869924* (*CLNK*), and (C) *rs2070025* (*FGA*). Overall effect is (A) OR 1.28, $P_{meta} = 0.03$; (B) OR = 1.37, $P_{meta} = 0.002$; and (C) OR = 1.34, $P_{meta} = 0.02$.

variants in *SLC2A9* (Figure 3), revealing *rs16869924* not to be in LD with the maximal gout-associated variant of *SLC2A9* in Polynesian populations (*rs3733591*, $r^2 = 0.05$). Linear regression analysis using a random effects model showed that *rs16869924* was not associated with serum urate concentrations ($\beta_{meta} 0.0055$ mmol/L, 95% CI -0.0063 to 0.0174 , $P = 0.36$). Similarly, neither *rs1635712* ($\beta_{meta} 0.0098$ mmol/L, 95% CI -0.0057 to 0.0253 , $P = 0.13$) nor *rs2070025* ($\beta_{meta} -0.0024$ mmol/L, 95% CI -0.0181 to 0.0134) were significantly associated with serum urate levels. Further, we included genotype at the lead gout-associated variant *rs3733591* (Figure 3; previously reported as associated with gout in Polynesian³⁶ and other⁸ populations) as a covariate in the logistic regression model testing for association with gout. Even with this adjustment, *rs16869924* remained associated with gout (OR 1.37, $P = 1.5 \times 10^{-4}$), demonstrating that *rs16869924* is associated with gout independent of genotype at *SLC2A9*.

DISCUSSION

We carried out in silico resequencing using a targeted panel of candidate genes with subsequent analysis for association with gout in the Aotearoa New Zealand Polynesian population. The derived alleles of missense variants p.Gly243Asp (*rs1635712*, *KIAA0319*), p.Ser65Gly (*rs16869924*, *CLNK*), and p.Ile6Val (*rs2070025*, *FGA*) increased the risk of gout. The 3 variants

can be regarded as candidate etiological functional variants, requiring validation by replication, fine-mapping, and experimental evidence, although prior evidence (below) raises the likelihood of genetic variation in *CLNK* being causal of gout.

Previously, several articles have reported *CLNK* gene variants to be associated with gout or serum urate.^{8,37,38} In a Tibetan sample set including 315 people with gout there was nominal evidence for association with gout of intronic variants *rs10033825* and *rs17467273*,³⁷ and in a Chinese sample set including 145 people with gout there was nominal evidence for association with gout of intronic variants *rs2041215* and *rs1686947*.³⁸ In our study, the *CLNK* Ser65Gly-derived allele associated with increased risk of gout (OR_{meta} 1.37, $P_{meta} = 0.0017$) but not with serum urate concentrations ($\beta_{meta} 0.0055$, $P = 0.36$). The variant *rs16869924* (*CLNK*) was associated with gout independent of genotype at *SLC2A9* (Figure 3), demonstrating it to be an independent risk locus for gout at the *SLC2A9* locus. Combined with previous studies, our data support *CLNK* to be a likely causal gene for gout. Given that *rs16869924* was the maximally associated variant within the *CLNK* gene, this variant represents a candidate causal variant for gout. The *CLNK* gene is a member of the SLP76 (lymphocyte cytosolic protein 2) family of adaptors, which is expressed in several cell types, including T cells, natural killer cells, and mast cells. Its expression appears to be dependent on sustained exposure to cytokines such as IL-2 and IL-3, and it

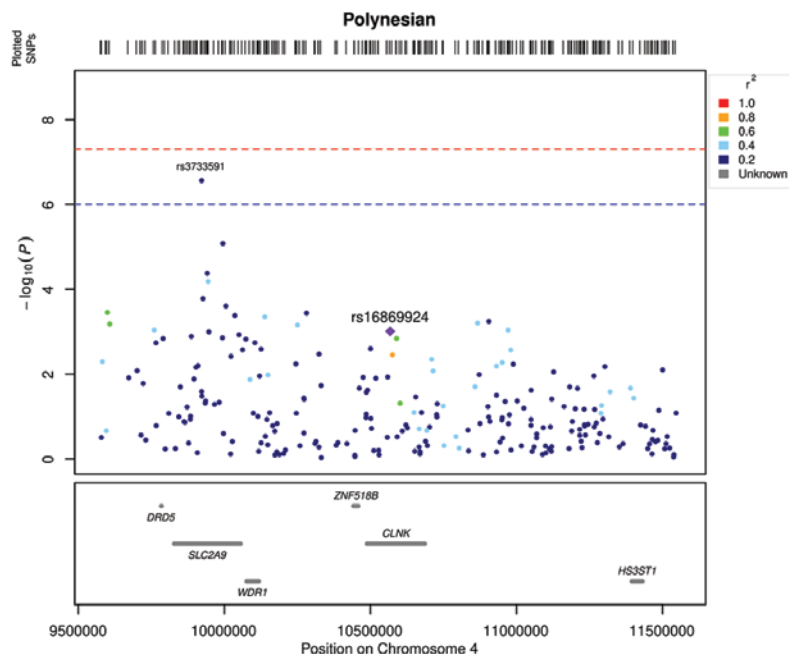


Figure 3. Regional association plot of *SLC2A9-CLNK*. The y-axis shows $-\log_{10} P$ values for the test of SNP association with gout in the genome-wide association study. LD (measured in r^2) with *rs16869924* was estimated based on the Polynesian data, and other SNPs are color-coded according to the extent of LD with *rs16869924*. The top SNP associated with gout in Polynesian populations at *SLC2A9* was *rs3733591* ($P = 2.75 \times 10^{-7}$, OR 0.64, 95% CI 0.54–0.77). The dotted red line is the genome-wide significance level ($P = 5 \times 10^{-8}$) and the dotted blue line the genome-wide suggestive significance level ($P = 1 \times 10^{-6}$). LD: linkage disequilibrium; SNP: single-nucleotide polymorphism.

plays a role in regulating immunoreceptor signaling.^{39,40} *CLNK* may directly or indirectly affect gout development through the signal transducer and activator of transcription pathway and the concentration of circulating immune complexes in the blood.^{41,42} Considering a role for *CLNK* in immune and inflammatory processes, as well as previous reports of association of the gene with gout, suggests *CLNK* to be of particular interest in gout pathogenesis in the New Zealand Polynesian population.

We cannot identify any previous studies where variants of *KIAA0319* and *FGA* are implicated in gout or hyperuricemia. In the Tin et al transancestral GWAS,⁸ there were no suggestive ($P < 1 \times 10^{-6}$) signals at the *KIAA0319* locus for either serum urate or gout. The *KIAA0319* gene has been associated with developmental dyslexia and is involved in neuronal migration.^{43,44} The deduced KIAA0319 protein contains several polycystic kidney disease domains, which may mediate the interaction between neurons and glial fibers during neuronal migration.⁴⁵ Mutations in the *FGA* (fibrinogen A α) gene lead to several disorders, including dysfibrinogenemia, hypofibrinogenemia, afibrinogenemia, and renal amyloidosis.^{46,47,48}

Of the nonassociated variants *ALDH16A1*, which codes for a member of the aldehyde dehydrogenase superfamily, deserves mention. *ALDH16A1* is enzymatically inactive but interacts with a number of proteins,^{49,50} including the γ subunit of AMPK (the gene encoding this protein [*PRKAG2*] maps within a locus associated with serum urate levels) and *GLUT4*, which belongs to the glucose transporter family to which *GLUT9* (encoded by

SLC2A9) belongs. Further, the derived G allele at p.Pro476Arg (*rs150414818*) has been predicted to inhibit the interaction between *ALDH16A1* and hypoxanthine-guanine phosphoribosyltransferase (HGPRT; encoded by *HPRT1*) that might contribute to hyperuricemia by disrupting the ability of HGPRT to salvage hypoxanthine from the urate-producing purine degradation pathway.⁵⁰ A previous study showed the Arg allele of *ALDH16A1* (*rs150414818*) to be a strong risk factor for gout (OR 3.12, $P = 1.5 \times 10^{-16}$) in the Icelandic population.⁵¹ In our study, the Polynesian common missense variant *rs78635115* (p.Ser178Pro allele of *ALDH16A1*) was not associated with gout (OR_{meta} 1.01, $P = 0.93$; Table 3).

In conclusion, this study identified 3 nominally gout-associated missense genetic variants in Aotearoa New Zealand Polynesian. Our data provide further support for *CLNK* as a risk gene at the *SLC2A9* locus, independent of *SLC2A9*.

ACKNOWLEDGMENT

The authors thank the many participants who so generously donated their genetic samples and other information for the study. The authors would also like to thank Jordyn Allan, Jill Drake, Roddi Laurence, Christopher Franklin, Meaghan House, Ria Akuhata-Brown, Nancy Aupouri, Carol Ford, and Gabrielle Sexton for recruitment, and NeSI (New Zealand eScience Infrastructure) for provision of compute resource.

REFERENCES

1. Dalbeth N, Dowell T, Gerard C, Gow P, Jackson G, Shuker C, et al.

- Gout in Aotearoa New Zealand: the equity crisis continues in plain sight. *N Z Med J* 2018;131:8-12.
2. Guillén AG, Te Karu L, Singh JA, Dalbeth N. Gender and ethnic inequities in gout burden and management. *Rheum Dis Clin North Am* 2020;46:693-703.
 3. Boocock J, Leask M, Okada Y, Matsuo H, Kawamura Y, Shi Y, et al; Asian Genetic Epidemiology Network (AGEN) Consortium. Genomic dissection of 43 serum urate-associated loci provides multiple insights into molecular mechanisms of urate control. *Hum Mol Genet* 2020;29:923-43.
 4. Li C, Li Z, Liu S, Wang C, Han L, Cui L, et al. Genome-wide association analysis identifies three new risk loci for gout arthritis in Han Chinese. *Nat Commun* 2015;6:7041.
 5. Matsuo H, Yamamoto K, Nakaoka H, Nakayama A, Sakiyama M, Chiba T, et al. Genome-wide association study of clinically defined gout identifies multiple risk loci and its association with clinical subtypes. *Ann Rheum Dis* 2016;75:652-9.
 6. Okada Y, Sim X, Go MJ, Wu JY, Gu D, Takeuchi F, et al. Meta-analysis identifies multiple loci associated with kidney function-related traits in East Asian populations. *Nat Genet* 2012;44:904-9.
 7. Phipps-Green AJ, Merriman ME, Topless R, Altaf S, Montgomery GW, Franklin C, et al. Twenty-eight loci that influence serum urate levels: analysis of association with gout. *Ann Rheum Dis* 2016;75:124-30.
 8. Tin A, Marten J, Halperin Kuhns VL, Li Y, Wuttke M, Kirsten H, et al. Target genes, variants, tissues and transcriptional pathways influencing human serum urate levels. *Nat Genet* 2019;51:1459-74.
 9. Major TJ, Dalbeth N, Stahl EA, Merriman TR. An update on the genetics of hyperuricaemia and gout. *Nat Rev Rheumatol* 2018;14:341-53.
 10. Dalbeth N, Phipps-Green A, Frampton C, Neogi T, Taylor WJ, Merriman TR. Relationship between serum urate concentration and clinically evident incident gout: An individual participant data analysis. *Ann Rheum Dis* 2018;77:1048-52.
 11. Chang WC, Jan Wu YJ, Chung WH, Lee YS, Chin SW, Chen TJ, et al. Genetic variants of PPAR-gamma coactivator 1B augment NLRP3-mediated inflammation in gouty arthritis. *Rheumatology* 2017;56:457-66.
 12. Chen Y, Ren X, Li C, Xing S, Fu Z, Yuan Y, et al. CARD8 rs2043211 polymorphism is associated with gout in a Chinese male population. *Cell Physiol Biochem* 2015;35:1394-400.
 13. Dong Z, Zhou J, Jiang S, Li Y, Zhao D, Yang C, et al. Effects of multiple genetic loci on the pathogenesis from serum urate to gout. *Sci Rep* 2017;7:43614.
 14. McKinney C, Stamp LK, Dalbeth N, Topless RK, Day RO, Kannangara DR, et al. Multiplicative interaction of functional inflammasome genetic variants in determining the risk of gout. *Arthritis Res Ther* 2015;17:288.
 15. Merriman TR, Terkeltaub R. PPARGC1B: Insight into the expression of the gouty inflammation phenotype: PPARGC1B and gouty inflammation. *Rheumatology* 2017;56:323-5.
 16. Rasheed H, Phipps-Green AJ, Topless R, Smith MD, Hill C, Lester S, et al. Replication of association of the apolipoprotein A1-C3-A4 gene cluster with the risk of gout. *Rheumatology* 2016;55:1421-30.
 17. Matsuo H, Ichida K, Takada T, Nakayama A, Nakashima H, Nakamura T, et al. Common dysfunctional variants in ABCG2 are a major cause of early-onset gout. *Sci Rep* 2013;3:2014.
 18. Tanner C, Boocock J, Stahl EA, Dobbyn A, Mandal AK, Cadzow M, et al. Population-specific resequencing associates the ATP-binding cassette subfamily C member 4 gene with gout in New Zealand Māori and Pacific men. *Arthritis Rheumatol* 2017;69:1461-9.
 19. Klück V, van Deuren RC, Cavalli G, Shaikat A, Arts P, Cleophas M, et al. Rare genetic variants in interleukin-37 link this anti-inflammatory cytokine to the pathogenesis and treatment of gout. *Ann Rheum Dis* 2020;79:536-44.
 20. Gosling AL, Boocock J, Dalbeth N, Harré Hindmarsh J, Stamp LK, Stahl EA, et al. Mitochondrial genetic variation and gout in Māori and Pacific people living in Aotearoa New Zealand. *Ann Rheum Dis* 2018;77:571-8.
 21. Krishnan M, Major TJ, Topless RK, Dewes O, Yu L, Thompson JMD, et al. Discordant association of the CREBRF rs373863828 A allele with increased BMI and protection from type 2 diabetes in Māori and Pacific (Polynesian) people living in Aotearoa/New Zealand. *Diabetologia* 2018;61:1603-13.
 22. Moors J, Leask M, Dalbeth N, Stamp LK, Harré-Hindmarsh J, Merriman TR, editors. Pacific-specific CETP genetic variant associates with increased HDL and low LDL in Polynesian populations [abstract]. *Diabetologia* 2019;S1:1149.
 23. Wallace SL, Robinson H, Masi AT, Decker JL, Mccarty DJ, Yü TF. Preliminary criteria for the classification of the acute arthritis of primary gout. *Arthritis Rheum* 1977;20:895-900.
 24. Stamp LK, Merriman TR, Singh JA. Expert opinion on emerging urate-lowering therapies. *Expert Opin Emerg Drugs* 2018;23:201-9.
 25. Wood R, Fermer S, Ramachandran S, Baumgartner S, Morlock R. Gout patients treated with conventional urate-lowering therapy: association between disease control and health-related quality of life. *J Rheumatol* 2016;43:1897-903.
 26. Phipps-Green AJ, Hollis-Moffatt JE, Dalbeth N, Merriman ME, Topless R, Gow PJ, et al. A strong role for the ABCG2 gene in susceptibility to gout in New Zealand Pacific island and Caucasian, but not Māori, case and control sample sets. *Hum Mol Genet* 2010;19:4813-9.
 27. Hahsler M, Piekenbrock M, Doran D. dbSCAN: fast density-based clustering with R. *J Stat Softw* 2019;91:1-30.
 28. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;43:491-8.
 29. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010;26:589-95.
 30. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297-303.
 31. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38:e164.
 32. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015;4:7.
 33. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010;26:2190-1.
 34. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26:2336-7.
 35. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res* 2019;47:D886-D94.
 36. Hollis-Moffatt JE, Gow PJ, Harrison AA, Highton J, Jones PB, Stamp LK, et al. The SLC2A9 nonsynonymous Arg265His variant and gout: evidence for a population-specific effect on severity. *Arthritis Res Ther* 2011;13:R85.
 37. Lan B, Chen P, Jiri M, He N, Feng T, Liu K, et al. WDR1 and CLNK gene polymorphisms correlate with serum glucose and

- high-density lipoprotein levels in Tibetan gout patients. *Rheumatol Int* 2016;36:405-12.
38. Jin TB, Ren Y, Shi X, Jiri M, He N, Feng T, et al. Genetic variations in the CLNK gene and ZNF518B gene are associated with gout in case-control sample sets. *Rheumatol Int* 2015;35:1141-7.
 39. Wu JN, Koretzky GA. The SLP-76 family of adapter proteins. *Semin Immunol* 2004;16:379-93.
 40. Yu J, Riou C, Davidson D, Minhas R, Robson JD, Julius M, et al. Synergistic regulation of immunoreceptor signaling by SLP-76-related adaptor CLNK and serine/threonine protein kinase HPK-1. *Mol Cell Biol* 2001;21:6102-12.
 41. Wang X, Shaw S, Amiri F, Eaton DC, Marrero MB. Inhibition of the JAK/STAT signaling pathway prevents the high glucose-induced increase in TGF- β and fibronectin synthesis in mesangial cells. *Diabetes* 2002;51:3505-9.
 42. Marrero MB, Banes-Berclci AK, Stern DM, Eaton DC. Role of the JAK/STAT signaling pathway in diabetic nephropathy. *Am J Physiol Renal Physiol* 2006;290:F762-8.
 43. Cope N, Harold D, Hill G, Moskvina V, Stevenson J, Holmans P, et al. Strong evidence that KIAA0319 on chromosome 6p is a susceptibility gene for developmental dyslexia. *Am J Hum Genet* 2005;76:581-91.
 44. Paracchini S, Thomas A, Castro S, Lai C, Paramasivam M, Wang Y, et al. The chromosome 6p22 haplotype associated with dyslexia reduces the expression of KIAA0319, a novel gene involved in neuronal migration. *Hum Mol Genet* 2006;15:1659-66.
 45. Velayos-Baeza A, Toma C, Paracchini S, Monaco AP. The dyslexia-associated gene KIAA0319 encodes highly N- and O-glycosylated plasma membrane and secreted isoforms. *Hum Mol Genet* 2008;17:859-71.
 46. Rowczenio D, Stensland M, de Souza GA, Strøm EH, Gilbertson JA, Taylor G, et al. Renal amyloidosis associated with 5 novel variants in the fibrinogen A alpha chain protein. *Kidney Int Rep* 2017;2:461-9.
 47. Neerman-Arbez M, de Moerloose P. Mutations in the fibrinogen gene cluster accounting for congenital afibrinogenemia: an update and report of 10 novel mutations. *Human Mutat* 2007;28:540-53.
 48. Cao Z, Dong Y, Zeng J, Zhu H, Xie X, Liu J, et al. Whole-exome sequencing identified novel mutations in FGA and FGG genes in the patients with decreased fibrinogen. *Thromb Res* 2019;177:79-82.
 49. Vasiliou V, Sandoval M, Backos DS, Jackson BC, Chen Y, Reigan P, et al. ALDH16A1 is a novel non-catalytic enzyme that may be involved in the etiology of gout via protein-protein interactions with HPRT1. *Chem Biol Interact* 2013;202:22-31.
 50. Sulem P, Gudbjartsson DF, Walters GB, Helgadóttir HT, Helgason A, Gudjonsson SA, et al. Identification of low-frequency variants associated with gout and serum uric acid levels. *Nat Genet* 2011;43:1127-30.