Supplementary Material

SNP-based heritability estimates of gout and its subtypes determined by genome-wide association studies of clinically defined gout

Yu Toyoda, Masahiro Nakatochi, Akiyoshi Nakayama, *et al.*

Correspondence to:
Hirotaka Matsuo (hmatsuo.ndmc@gmail.com)

Supplementary Methods

Ethics Statement

Data and sample collection from the cohorts participating in the present study were approved by the individual research ethics committees (National Defense Medical College and Nagoya University). All the studies were performed according to the guidelines of the Declaration of Helsinki. All the participants had provided their written informed consent. This research was carried out without patient involvement.

Gout patients as cases and non-gout subjects for controls

An overview of the characteristics of the study populations is provided in **Supplementary Table S6**. The analyzed set consists of two groups of Japanese males: a clinically-defined gout group consisting of 3,053 patients and a non-gout (control) group consisting of 5,637 subjects, which are respectively the same cohort and a larger cohort than that with 4,554 subjects derived from our previous study, respectively [1]. In brief, patients with known clinical parameters were recruited from Japanese male outpatients at gout clinics as previously described [1]; all the patients were clinically diagnosed with primary gout according to the criteria established by the American College of Rheumatology [2]. Of the 3,053 gout cases, urinary data were available for 2,137 patients; based on clinical parameters, their subtypes of gout (**Fig. 1**) were diagnosed as previously described [1, 3], which resulted in dividing all gout cases into four subtypes: the renal underexcretion (RUE) type (654 cases), the renal overload (ROL) type (486 cases), the combined (RUE + ROL) type (905 cases), and the normal type (92 cases) (**Supplementary Table S7**). In addition to the previously assigned 4,554 subjects as normouricemia controls (serum urate ≤ 7 mg/dL; without past histories of gout) [1] who had taken part in the Japan Multi-Institutional Collaborative Cohort Study (J-MICC Study) [4, 5], in this study we enrolled 1,083 hyperuricemia subjects without gout as asymptomatic hyperuricemia controls (serum urate > 7 mg/dL; without gout), who were excluded in the previous study [1], to extend the range covered for serum urate concentrations in controls. With these subjects newly assigned in this study, we conducted single nucleotide polymorphism (SNP) genotyping for a genome-wide association study (GWAS). Meta-analyses using the obtained and previous data were then carried out as described below.

Genotyping, quality control, and genotype imputation

Information regarding study specific genotyping, imputation, and analysis tools is provided in **Supplementary Table S8**. Genotyping, quality control, and imputation for the Japonica Array data set that were obtained using a Japonica SNP Array [6] or the Illumina Array data set, obtained using HumanOmniExpress or HumanOmniExpress-Exome BeadChip Arrays (Illumina, San Diego, CA, USA) have been described previously [1].

Post-imputation quality control was conducted as reported previously [7], with minor modifications; SNPs with an imputation quality of $r^2 < 0.3$ or a minor allele frequency (MAF) of < 1% or 0.1% were excluded. To identify studies with inflated GWAS significance, which can result from population stratification, we computed the genomic control lambda [8] and the intercept of LD score regression [9]. We calculated the genomic control lambda using R statistical software (ver. 3.6) (http://www.r-project.org/).

After passing post-imputation quality control, with the Japonica Array data set, 1,028 case subjects and 1,125 control subjects as well as 8,581,683 (MAF ≥ 1%) or 10,779,400 (MAF ≥ 0.1%) SNPs remained for the subsequent GWAS analysis; with the Illumina Array data set, 2,025 case subjects and 4,512 control subjects as well as 8,359,615 (MAF ≥ 1%) or 10,473,049 (MAF ≥ 0.1%) SNPs remained for the subsequent GWAS analysis.

Association analysis for SNPs and clinically-defined gout including its subtypes

The association of SNPs with clinically-defined gout including its subtypes was assessed using logistic regression analysis (generalized linear model), as described previously [1] with minor modifications. In brief, the dependent variable was gout label (case = 1, control = 0), and the independent variables included imputed genotypes of each SNP and covariates—the first ten principal component scores. The effect sizes and standard errors estimated in logistic regression analysis were used in the subsequent meta-analysis. The association analysis was performed using the Efficient and Parallelizable Association Container Toolbox (https://genome.sph.umich.edu/wiki/EPACTS).

Meta-analyses

Meta-analyses were performed using a total of 3,053 cases (for all gout; numbers of cases for each subtype are described above) and 5,637 controls from the two data set as described previously [1]. In brief, the association results for each SNP across the studies were combined with METAL software [10] using the fixed-effects inverse-variance-weighted method. Heterogeneity of effect sizes was assessed via the I^2 index. The genome-wide significance level α was set to a P value of $< 5 \times 10^{-8}$. As an overview of the results, Manhattan plots of the genome-wide meta-analyses for genetic loci are provided in **Fig. 1B–E**.

Estimation of the SNP-based heritability of gout

Based on the obtained results of our Japanese meta-analyses, SNP-based heritability of gout was estimated using LD score regression [9] as described previously [7]. In brief, the heritability estimates were calculated from the summary statistics of 1,029,593 SNPs for all gout, which have MAF ≥ 1% and were not palindromic SNPs. With gout subtypes, three parts of the summary statistics of 1,029,179 (for the RUE type), 1,029,019 (for the ROL type), and 1,029,173 (for the combined type) SNPs were used, respectively. In a similar manner, single nucleotide variation (SNV)-based heritability estimates were calculated from 1,058,455 (for all gout), 1,058,284 (for RUE type), 1,058,247 (for ROL type), and 1,058,367 (for combined type) SNVs with MAF ≥ 0.1%.

Partitioning heritability

To partition heritability from summary statistics into multiple categories derived from functional annotations for certain parts of a genome, stratified LD score regression was conducted as described previously [11, 12]. In this study, we employed the full baseline model with 53 categories [11] and examined in which functional categories the heritability of gout is enriched. Significant heritability enrichments were defined when the false discovery rate (FDR) was below 0.1, as calculated using the Benjamini and Hochberg method [13].

Supplementary References

1 Nakayama A, Nakatochi M, Kawamura Y, *et al.* Subtype-specific gout susceptibility loci and enrichment of selection pressure on ABCG2 and ALDH2 identified by subtype genome-wide meta-analyses of clinically defined gout patients. Ann Rheum Dis 2020;79(5):657-65.

2 Wallace SL, Robinson H, Masi AT, Decker JL, McCarty DJ, Yu TF. Preliminary criteria for the classification of the acute arthritis of primary gout. Arthritis Rheum 1977;20(3):895-900.

3 Ichida K, Matsuo H, Takada T, *et al.* Decreased extra-renal urate excretion is a common cause of hyperuricemia. Nat Commun 2012;3:764.

4 Asai Y, Naito M, Suzuki M, *et al.* Baseline data of Shizuoka area in the Japan Multi-Institutional Collaborative Cohort Study (J-MICC Study). Nagoya J Med Sci 2009;71(3-4):137-44.

5 Hamajima N, J-Micc Study Group. The Japan Multi-Institutional Collaborative Cohort Study (J-MICC Study) to detect gene-environment interactions for cancer. Asian Pac J Cancer Prev 2007;8(2):317-23.

6 Kawai Y, Mimori T, Kojima K, *et al.* Japonica array: improved genotype imputation by designing a population-specific SNP array with 1070 Japanese individuals. J Hum Genet 2015;60(10):581-7.

7 Nakatochi M, Kanai M, Nakayama A, *et al.* Genome-wide meta-analysis identifies multiple novel loci associated with serum uric acid levels in Japanese individuals. Commun Biol 2019;2:115.

8 Devlin B, Roeder K. Genomic control for association studies. Biometrics 1999;55(4):997-1004.

9 Bulik-Sullivan BK, Loh PR, Finucane HK, *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 2015;47(3):291-5.

10 Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26(17):2190-1.

11 Finucane HK, Bulik-Sullivan B, Gusev A, *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. Nat Genet 2015;47(11):1228-35.

12 Kanai M, Akiyama M, Takahashi A, *et al.* Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. Nat Genet 2018;50(3):390-400.

13 Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol 1995;57(1):289-300.

14 Matsuo H, Yamamoto K, Nakaoka H, *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(4):652-9.

Supplementary Table S1. Estimation of the SNP-based heritability of gout and its subtypes using clinically-defined gout patients in Japanese populations

Trait	Prevalence*	*vs.* non-gout		*vs.* normouricemia	
		h^2_g	SE	h^2_g	SE
All gout	0.0106	0.279	0.085	0.333	0.098
RUE type	0.0032	0.332	0.118	0.342	0.120
ROL type	0.0024	0.355	0.135	0.397	0.148
Combined type	0.0045	0.309	0.099	0.368	0.110

*Data are derived from previous studies [1, 14]. SNPs with minor allele frequency ≥ 1% were used for calculation. SNP, single nucleotide polymorphism; RUE, renal underexcretion; ROL, renal overload; h^2_g, heritability; SE, standard error of h^2.

Supplementary Table S2. Genetic correlations based on SNPs among major subtypes of gout

Trait 1	Trait 2	*N*	r_g	SE	*P*
Combined type	RUE type	1,039,419	0.683	0.193	0.4×10^{-4}
Combined type	ROL type	1,039,330	0.623	0.204	2.2×10^{-4}
ROL type	RUE type	1,039,423	0.737	0.189	9.9×10^{-5}

RUE, renal underexcretion; ROL, renal overload; r_g, genetic correlation coefficient; SE, standard error of r_g.

Supplementary Table S3. Statistically significant heritability enrichments for clinically-defined gout (all gout)

Categories*	Prop. SNVs	Prop. h^2_g	Enrichment	Enrichment *P*-value	Q-value
Enhancer + 500bp	0.089	0.547 (0.133)	6.115 (1.486)	1.12×10^{-3}	0.019
Repressed + 500bp	0.720	0.279 (0.120)	0.387 (0.166)	0.75×10^{-3}	0.019
Super Enhancer + 500bp	0.170	0.440 (0.085)	2.581 (0.500)	0.66×10^{-3}	0.019
H3K4me1	0.423	1.126 (0.294)	2.663 (0.695)	6.19×10^{-3}	0.081

Each value for standard error is noted in brackets. *Functional categories are from Finucane *et al.* [11]; categories with false discovery rate (Q-value) < 0.1 were defined as statistically significant in this analysis.
H3k4me1, monomethylation of histone H3 at lysine 4; Prop., proportion; SNVs, single nucleotide variations with minor allele frequency ≥ 0.1%; h^2_g, heritability.

Supplementary Table S4. Estimation of the SNV-based heritability of gout and its subtypes using clinically-defined gout patients in Japanese populations

Trait	Prevalence*	*vs.* non-gout		*vs.* normouricemia	
		h^2_g	SE	h^2_g	SE
All gout	0.0106	0.294	0.083	0.346	0.096
RUE type	0.0032	0.349	0.118	0.362	0.120
ROL type	0.0024	0.379	0.131	0.428	0.144
Combined type	0.0045	0.334	0.098	0.395	0.110

*Data are derived from previous studies [1, 14]. SNVs with minor allele frequency ≥ 0.1% were used for calculation. SNV, single nucleotide variation; RUE, renal underexcretion; ROL, renal overload; h^2_g, heritability; SE, standard error of h^2_g.

Supplementary Table S5. Significant loci identified in the present genome-wide meta-analysis of SNPs with minor allele frequency ≥ 1%

SNP*	Locus	Position†	Gene	Alleles		Illumina array				Japonica array				Meta-analysis			
						RAF				RAF							
				Risk	Non-risk	Case	Ctrl	OR (95%CI)	*P* value	Case	Ctrl	OR (95%CI)	*P* value	OR (95%CI)	*P* value	I^2	Het*P*
All gout patients																	
rs6547692	2p23.3	27734972	*GCKR*	G	A	0.623	0.547	1.30 (1.20–1.41)	7.10×10^{-11}	0.608	0.563	1.18 (1.04–1.35)	1.07×10^{-2}	1.27 (1.18–1.36)	5.55×10^{-12}	34.3	0.22
rs3775948	4p16.1	9995182	*SLC2A9*	C	G	0.680	0.587	1.55 (1.42–1.68)	1.16×10^{-24}	0.673	0.565	1.61 (1.41–1.84)	1.47×10^{-12}	1.56 (1.46–1.68)	1.44×10^{-35}	0	0.62
rs4148155	4q22.1	89054667	*ABCG2*	G	A	0.454	0.289	2.04 (1.87–2.21)	2.80×10^{-63}	0.462	0.281	2.17 (1.90–2.48)	9.86×10^{-30}	2.07 (1.93–2.23)	4.49×10^{-91}	0	0.43
rs1184804	6p22.2	25868226	*SLC17A3*	C	T	0.874	0.829	1.31 (1.16–1.48)	2.00×10^{-5}	0.868	0.832	1.45 (1.21–1.75)	7.28×10^{-5}	1.35 (1.22–1.49)	9.00×10^{-9}	0	0.35
rs3129473	10q23.2	88922846	*FAM35A*	G	C	0.424	0.367	1.37 (1.25–1.49)	9.21×10^{-13}	0.427	0.379	1.29 (1.12–1.48)	2.94×10^{-4}	1.35 (1.25–1.45)	1.48×10^{-15}	0	0.50
rs145954970	11q13.1	64273830	*LOC100996455–SLC22A11*	C	G	0.995	0.976	7.93 (4.46–14.08)	1.71×10^{-12}	0.997	0.971	23.47 (7.25–76.01)	1.41×10^{-7}	9.77 (5.83–16.38)	5.05×10^{-18}	62.2	0.10
rs671	12q24.12	112241766	*ALDH2*	G	A	0.823	0.737	1.80 (1.63–1.99)	5.64×10^{-32}	0.821	0.687	2.00 (1.73–2.32)	4.96×10^{-20}	1.86 (1.72–2.02)	4.97×10^{-50}	23.7	0.25
rs76499759	13q22.1	73568511	*PIBF1*	A	G	0.219	0.181	1.29 (1.17–1.42)	2.34×10^{-7}	0.212	0.183	1.21 (1.04–1.41)	1.61×10^{-2}	1.27 (1.17–1.38)	1.51×10^{-8}	0	0.49
rs7224656	17q23.2	59434815	*BCAS3*	A	T	0.447	0.404	1.19 (1.10–1.30)	3.47×10^{-5}	0.462	0.409	1.28 (1.13–1.47)	1.94×10^{-4}	1.22 (1.14–1.31)	3.95×10^{-8}	0	0.35
RUE type gout patients																	
rs3775948	4p16.1	9995182	*SLC2A9*	C	G	0.705	0.587	1.75 (1.50–2.04)	1.47×10^{-12}	0.717	0.565	1.98 (1.56–2.50)	1.45×10^{-8}	1.81 (1.59–2.06)	1.79×10^{-19}	0	0.39
rs4148155	4q22.1	89054667	*ABCG2*	G	A	0.388	0.289	1.56 (1.35–1.81)	2.47×10^{-9}	0.382	0.281	1.57 (1.26–1.95)	6.82×10^{-5}	1.57 (1.38–1.77)	7.46×10^{-13}	0	0.99
rs9420434	10q23.2	88843209	*GLUD1(SHLD2)*	C	T	0.316	0.253	1.45 (1.25–1.70)	1.79×10^{-6}	0.342	0.247	1.61 (1.28–2.03)	4.60×10^{-5}	1.50 (1.32–1.71)	4.52×10^{-10}	0	0.46
rs76741582	11q13.1	64247850	*LOC100996455–SLC22A11*	T	C	0.028	0.010	3.05 (1.89–4.94)	5.49×10^{-6}	0.036	0.008	5.23 (2.51–10.90)	9.83×10^{-6}	3.59 (2.40–5.37)	4.81×10^{-10}	31.1	0.23
rs4646776	12q24.12	112230019	*ALDH2*	G	C	0.819	0.734	1.79 (1.50–2.15)	3.28×10^{-10}	0.796	0.683	1.78 (1.38–2.32)	1.33×10^{-5}	1.79 (1.54–2.08)	2.07×10^{-14}	0	0.97
ROL type gout patients																	
rs4148155	4q22.1	89054667	*ABCG2*	G	A	0.515	0.289	2.67 (2.26–3.14)	9.99×10^{-32}	0.493	0.281	2.34 (1.81–3.03)	1.40×10^{-10}	2.57 (2.24–2.95)	1.43×10^{-40}	0	0.40
rs12231737	12q24.13	112574616	*TRAFD1(ALDH2)*	C	T	0.812	0.757	1.88 (1.46–2.43)	1.11×10^{-6}	0.810	0.668	2.10 (1.52–2.91)	7.95×10^{-6}	1.96 (1.61–2.40)	4.46×10^{-11}	0	0.60
Combined type gout patients																	
rs1260326	2p23.3	27730940	*GCKR*	T	C	0.647	0.550	1.42 (1.25–1.62)	1.95×10^{-7}	0.615	0.564	1.23 (1.02–1.47)	2.65×10^{-2}	1.35 (1.21–1.50)	3.58×10^{-8}	39.1	0.20
rs3775948	4p16.1	9995182	*SLC2A9*	C	G	0.690	0.587	1.61 (1.40–1.85)	2.14×10^{-11}	0.672	0.565	1.57 (1.3–1.88)	2.16×10^{-6}	1.59 (1.42–1.78)	2.41×10^{-16}	0	0.83
rs2231142	4q22.1	89052323	*ABCG2*	T	G	0.474	0.289	2.27 (1.98–2.59)	1.06×10^{-32}	0.474	0.281	2.30 (1.91–2.76)	5.37×10^{-19}	2.28 (2.04–2.54)	5.20×10^{-50}	0	0.90
rs3129399	10q23.2	88951662	*FAM35A*	G	A	0.458	0.377	1.56 (1.35–1.80)	1.07×10^{-9}	0.438	0.399	1.30 (1.06–1.61)	1.37×10^{-2}	1.47 (1.31–1.66)	1.33×10^{-10}	49.8	0.16
rs56093838	11q13.1	64383327	*NRXN2*	C	T	0.489	0.428	1.36 (1.18–1.55)	1.27×10^{-5}	0.493	0.409	1.48 (1.22–1.78)	4.88×10^{-5}	1.40 (1.25–1.56)	3.28×10^{-9}	0	0.46
rs116873087	12q24.13	112511913	*NAA25(ALDH2)*	G	C	0.828	0.759	2.12 (1.72–2.60)	1.12×10^{-12}	0.845	0.686	2.50 (1.97–3.17)	4.19×10^{-14}	2.27 (1.95–2.66)	5.31×10^{-25}	4.9	0.31

*dbSNP rs number, †SNP positions are based on NCBI human genome reference sequence Build hg19. SNP, single nucleotide polymorphism; RAF, risk allele frequency; Ctrl, control; OR, odds ratio; CI, confidence interval; Het*P*, heterogeneity *P* value; RUE, renal underexcretion; ROL, renal overload.

Supplementary Table S6. Characteristics of Japanese males without medication for urate-lowering therapy who took part in the study

	Japonica array platform		Illumina array platform		Total	
	Case	Control	Case	Control	Case	Control
Number	1,028	1,125	2,025	4,512	3,053	5,637
Age (year)	44.6 ± 11.4	53.9 ± 9.2	48.0 ± 11.8	55.3 ± 9.3	46.9 ± 11.8	55.0 ± 9.3
Body-mass index (kg/m^2)	25.2 ± 3.6	23.1 ± 2.9	25.0 ± 3.5	23.8 ± 3.1	25.0 ± 3.6	23.7 ± 3.1

Data are expressed as means ± SD; case data are from Nakayama *et al.* [1], under the terms of the Creative Commons Attribution 4.0 international License (https://creativecommons.org/licenses/by/4.0/.).

Supplementary Table S7. Clinical parameters of subtype gout cases

		Number	Age (year)	BMI (kg/m^2)	SUA (mg/dL)	FE_{UA} (%)	UUE (mg/h/1.73m^2)
Japonica array platform	RUE type	211	47.7 ± 10.8	24.9 ± 3.4	8.36 ± 1.28	3.78 ± 0.82	20.6 ± 3.6
	ROL type	137	48.7 ± 11.2	24.3 ± 2.9	8.17 ± 1.18	6.58 ± 0.96	37.6 ± 8.6
	Combined type	349	43.6 ± 10.0	25.4 ± 3.7	8.75 ± 1.22	4.34 ± 0.69	31.9 ± 7.3
Illuimina array platform	RUE type	443	47.6 ± 11.0	25.0 ± 3.8	8.38 ± 1.15	3.78 ± 0.82	20.2 ± 3.5
	ROL type	349	50.0 ± 10.9	24.5 ± 3.0	8.20 ± 1.14	6.89 ± 1.62	36.4 ± 9.2
	Combined type	556	44.4 ± 9.9	25.9 ± 3.8	8.76 ± 1.24	4.41 ± 0.72	31.6 ± 6.5
Total	RUE type	654	47.6 ± 10.9	24.9 ± 3.6	8.37 ± 1.19	3.78 ± 0.82	20.3 ± 3.5
	ROL type	486	49.7 ± 11.0	24.4 ± 2.9	8.19 ± 1.15	6.81 ± 1.47	36.7 ± 9.0
	Combined type	905	44.1 ± 9.9	25.6 ± 3.7	8.75 ± 1.22	4.39 ± 0.71	31.7 ± 6.8

Data are expressed as means ± SD; some case data are from Nakayama *et al.* [1], under the terms of the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/.).

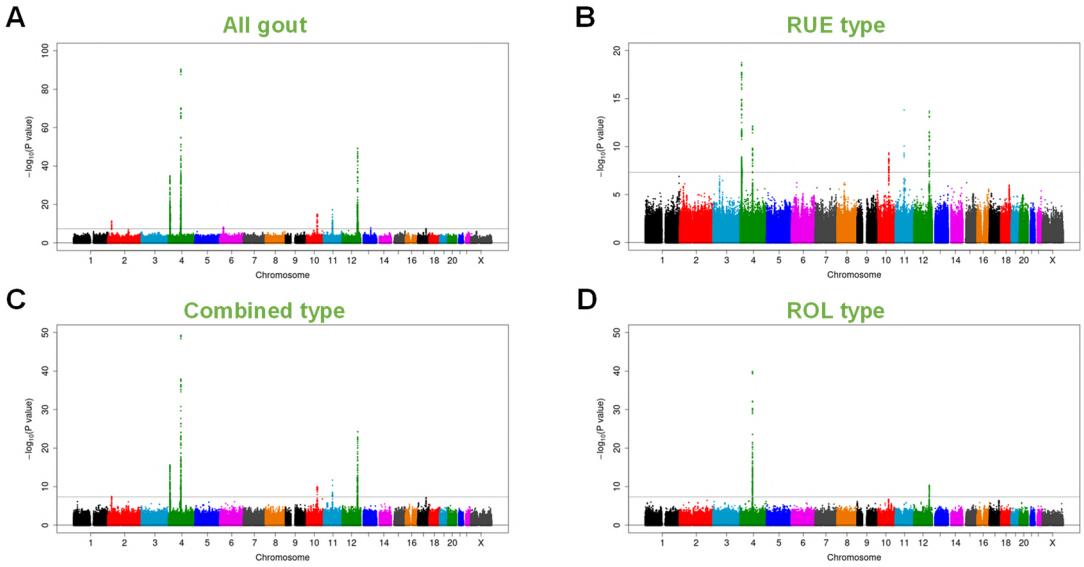
Normal type, which is characterized by UUE ≤ 25 mg/h/1.73 m^2 and FE_{UA} ≥ 5.5%, was excluded in this study due to the small number of patients (total 92 cases). BMI, body-mass index; SUA, serum uric acid; FE_{UA}, fractional excretion of uric acid; UUE, urinary urate excretion; RUE, renal underexcretion; ROL, renal overload.

Supplementary Table S8. Genotyping, imputation, and association testing

Study name	Genotyping platform	Genotype calling algorithm	Pre-imputation QC					#SNPs for imputation	Imputation reference	Imputation software	Post-imputation QC		Association study software	Study specific covariates for association study
			Sample call rate cutoff	Other criteria	SNP call rate cutoff	SNP MAF cutoff	SNP HWE P cutoff				SNP r^2 cutoff	SNP MAF cutoff		
Japonica array platform	Japonica SNP Array	Genome Studio	0.98	Sex mismatches; related samples (IBD 0.1875); samples not mapping to JPT (1000 genomes phase 3)	0.98	0.01	1×10^{-6}	603,009	1000G phase3v5; individuals of all ancestries	SHAPEIT2 minimac3	0.3	0.001 or 0.01	EPACTS	Top 10 principal components
Illumina array platform	Illumina HumanOmni Express or HumanOmni ExpressExome	Genome Studio	0.98	Sex mismatches; related samples (IBD 0.1875); samples not mapping to JPT (1000 genomes phase 3)	0.98	0.01	1×10^{-6}	553,321	1000G phase3v5; individuals of all ancestries	SHAPEIT2 minimac3	0.3	0.001 or 0.01	EPACTS	Top 10 principal components

QC, quality control; MAF, minor allele frequency; HWE, Hardy-Weinberg equilibrium; #SNPs, the number of autosomal single nucleotide polymorphisms.

Supplementary Figure



Supplementary Figure S1. Manhattan plots of genome-wide meta-analyses of gout using single nucleotide variations (SNVs) with minor allele frequency (MAF) ≥ 0.1%

For all gout, (*A*); RUE type, (*B*); combined (RUE +ROL) type, (*C*); ROL type, (*D*). The horizontal axis represents chromosomal positions and the vertical axis indicates the $-\log_{10}(P \text{ value})$ for the assessment of the association. Horizontal lines represent the genome-wide significance threshold ($\alpha = 5 \times 10^{-8}$). While all data shown in **Fig. 1** were derived from SNPs with MAF ≥ 1%, all data in this **Supplementary Figure S1** were from SNVs with MAF ≥ 0.1%.