Replication of Gout/Urata Concentrations GWAS Susceptibility Loci Associated with Gout in a Han Chinese Population

Zhiqiang Li1,2,3, Zhaowei Zhou1, Xu Hou3,4, Dajiang Lu5, Xuan Yuan2,4, Jie Lu4,6, Can Wang2,4, Lin Han2,4, Lingling Cui3,4, Zhen Liu2,4, Jianhua Chen3, Xiaoyu Cheng3,4, Keke Zhang2,4, Xuefeng Wang2,4, Xinde Li2,4, Qing-Sheng Mi6,7,8, Yongyong Shi1,2,3 & Changgui Li1,2,4

Gout is a chronic disease resulting from elevated serum urate (SU). Previous genome-wide association studies (GWAS) have identified dozens of susceptibility loci for SU/gout, but few have been conducted for Chinese descent. Here, we try to extensively investigate whether these loci contribute to gout risk in Han Chinese. A total of 2255 variants in linkage disequilibrium (LD) with GWAS identified SU/gout associated variants were analyzed in a Han Chinese cohort of 1255 gout patients and 1848 controls. Cumulative genetic risk score analysis was performed to assess the cumulative effect of multiple “risk” variants on gout incidence. 23 variants (41%) of LD pruned variants set (n = 56) showed nominal association with gout in our sample (p < 0.05). Some of the previously reported gout associated loci (except ALDH16A1), including ABCG2, SLC2A9, GCKR, ALDH2 and CNH2, were replicated. Cumulative genetic risk score analyses showed that the risk of gout increased for individuals with the growing number (≥8) of the risk alleles on gout associated loci. Most of the gout associated loci identified in previous GWAS were confirmed in an independent Chinese cohort, and the SU associated loci also confer susceptibility to gout. These findings provide important information of the genetic association of gout.

Gout is characteristic of acute arthritis, joint deformity and severe pain caused by deposition of monosodium urate crystals in and around synovial tissue. Elevated serum urate (SU) levels are the most important risk factor of gout. Gout is a multifactorial disease dominated by genetic component with heritability of SU estimated to be 73%1. In the recent decade, much effort especially genome-wide association studies (GWAS) have attempted to clarify such contribution and identified dozens of susceptibility loci for SU/gout2–15. These studies from different populations (European, African Americans, Japanese and Chinese) provided important clues for better understanding the etiology of gout and some evidences of heterogeneity across populations2–14. To our knowledge, only two GWAS, which were published very recently, were conducted for clinically defined gout cases only. One was performed in the Han Chinese population13, and the other in the Japanese population 14. Most of other studies for SU levels (and gout) have primarily been conducted in populations of European descent2–12, 15. It is of great interest to replicate the candidate loci for European and/or other populations in Han Chinese population.

1The Biomedical Sciences Institute of Qingdao University, Qingdao Branch of SJTU Bio-X Institutes & the Affiliated Hospital of Qingdao University, Qingdao, 266003, P.R. China. 2Shandong Provincial Key Laboratory of Metabolic Disease, the Affiliated Hospital of Qingdao University & the Metabolic Disease Institute of Qingdao University, Qingdao, 266003, P.R. China. 3Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders (Ministry of Education) and the Collaborative Innovation Center for Brain Science, Shanghai Jiao Tong University, Shanghai, 200030, P.R. China. 4Shandong Gout Clinical Medical Center, Qingdao, 266003, P.R. China. 5School of Kinesiology, Shanghai University of Sport, Shanghai, 200048, P.R. China. 6Henry Ford Immunology Program, Henry Ford Health System, 1 Ford Place, Detroit, MI, 48202, USA. 7Department of Dermatology, Henry Ford Health System, 1 Ford Place, Detroit, MI, 48202, USA. 8Department of Internal Medicine, Henry Ford Health System, 1 Ford Place, Detroit, MI, 48202, USA. Zhiqiang Li, Zhaowei Zhou, Xu Hou and Dajiang Lu Contributed equally to this work. Correspondence and requests for materials should be addressed to Y.S. (email: shiyongyong@gmail.com) or C.L. (email: lichanggui@medmail.com.cn)
genetic studies for SU/gout have been conducted in Han Chinese\textsuperscript{16–19}. However, most of them examined only a small minority of loci. In the present study, we try to determine whether the previously identified SU/gout loci affect susceptibility to gout in Chinese using our recent gout GWAS dataset.

Methods

Samples, genotyping and variants selection. All samples including 1255 clinically ascertained gout patients and 1848 healthy controls were of Han Chinese males and signed written informed consent, as described in our recent gout GWAS paper (Supplementary Methods)\textsuperscript{13}. Clinical characteristics for the samples were shown in Supplementary Table S1. Genotyping was conducted using Affymetrix Axiom Genome-Wide CHB Array. Detailed methods of quality control and imputation were described as done previously\textsuperscript{13}, and a brief description was shown in Supplementary Methods. All the previously identified genome-wide significant loci (p < 5.0 × 10\textsuperscript{-8}) related to gout/SU were obtained from the NHGRI GWAS catalog (as to May 12, 2013) and further fine-mapping or mutation analysis studies\textsuperscript{20–22} (Supplementary Tables S2 and S3). Considering the linkage disequilibrium (LD) patterns might differ across different ethnicities for the same susceptibility locus, we included all the available variants those are in LD (r\textsuperscript{2} > 0.6) with the genome-wide significant variants based on 1000 Genomes Project datasets. A total of 2255 variants for 1255 gout patients and 1848 controls were kept for subsequent analyses.

Statistical analysis. Association analysis was performed using the logistic regression with 20 principal components as covariates for correcting the potential population stratification (PCA adjustment analysis, Supplementary Methods). In order to approximate the number of independent variants within each region, we pruned the variants based on LD. A total of 56 LD pruned variants were generated using a r\textsuperscript{2} threshold of 0.2. The simple Bonferroni correction for multiple comparisons (n = 2255) was applied, thus 2.22 × 10\textsuperscript{-5} (0.05/2255) was set as the statistical significance level. An uncorrected p value of 0.05 was considered as nominal evidence for association. For the variants in the gout associated loci, an evidence of nominal association was treated as a successful replication, considering pervious evidences for the associations between these loci and gout were solid.

The association and LD prune analyses were performed using PLINK\textsuperscript{23}. The exact binomial test was performed using R package, by comparing the direction of effect sizes of the tested SNPs between our dataset and the previous reports. The p value was generated under the null hypothesis (H0: p = 0.50). Cumulative genetic risk score analysis was conducted by counting risk alleles in an unweighted method for each individual and calculating the effect on gout risk using logistic regression analysis adjusting for the covariates of principal components.

The study protocol was approved by the Ethics Committee of the Affiliated Hospital, Qingdao University. All procedures were conducted in accordance with the Declaration of Helsinki\textsuperscript{24}

Data availability. The results are available upon request by contacting Li CG or Shi YY. Any additional data (beyond those included in the main text and Supplementary Information) that support the findings of this study are also available from the corresponding author upon request.

Ethics approval. This study was approved by the relevant ethics review board at the Affiliated Hospital of Qingdao University.

Results

Based on the LD data of the 1000 Genomes Project datasets for different continental populations (Mixed American, East Asian and European), a total of 2255 variants, which are in LD (r\textsuperscript{2} > 0.6) with the previous identified genome-wide significant variants for gout/SU, are tested in this study. After LD (r\textsuperscript{2} < 0.2) pruning, 56 LD independent variants were generated. Of the 56 LD independent variants, 23 variants (41%) showed association with gout at p < 0.05 in our sample (Supplementary Table S4). And 11 of the significant variants were from the gout associated loci (GCKR, ABCG2, ABC2, SLC2A9, CNIH2, MYL2-CUX2 (ALDH2) and BCA3). The strongest association signal was observed in the ABCG2 locus (rs1481012, p = 8.96 × 10\textsuperscript{-11}, OR = 1.890, Table 1), which is consistent with Köttgen et al’s report (rs1481012, p = 2.00 × 10\textsuperscript{-32}, OR = 1.730)\textsuperscript{11,21}. The top significant association was observed at rs11722228 (p = 2.40 × 10\textsuperscript{-6}, OR = 1.619) for the SLC2A9 locus (Table 1). These two variants survived Bonferroni correction for multiple testing (p < 2.22 × 10\textsuperscript{-5}). The most highly associated SNP at the GCKR locus is rs6547692 (p = 2.20 × 10\textsuperscript{-4}, OR = 0.696). The CNIH2 and MYL2-CUX2 (ALDH2) loci are two novel gout loci identified in recent Japanese studies\textsuperscript{14,22}. We observed nominal associations at rs801733 (CNIH2, p = 0.026, OR = 0.428) and rs11066008 (MYL2-CUX2 (ALDH2), p = 2.94 × 10\textsuperscript{-5}, OR = 0.666) (Table 1), which are in strong LD with the previously identified gout associated variants (rs4073582 and rs671, r\textsuperscript{2} = 0.96 and 0.79, respectively), and the directions of effects for both variants were consistent with the previous reports\textsuperscript{14,22}. The BCA3 locus is one of the novel gout loci identified in our previous report\textsuperscript{3}. In addition to this locus, our previous study identified another two gout loci at RFX3 (rs12236871) and KCNQ1 (rs179785) (Supplementary Table S3)\textsuperscript{3}. Taken together, all gout associated loci that reached the genome-wide significance level in the previous GWAS reports were replicated, except ALDH16A1.

The previously identified susceptibility SNPs were usually considered as more important variants, especially the non-synonymous ones should be given priorities. Because these variants are most likely to have functional effect on gout risk using logistic regression analysis adjusting for the covariates of principal components.

The previously reported non-synonymous variants were available in our dataset (Table 2), and all the gout-risk and SU-raising alleles were overrepresented in our cases (Exact binominal test p = 7.81 × 10\textsuperscript{-3}). Of them, two variants exhibited statistically significant associations (ABCG2 Q141K (rs2231142), p = 3.83 × 10\textsuperscript{-10} and SLC17A1 I269T (rs1165196), p = 1.94 × 10\textsuperscript{-10}) and three showed nominal significant associations (SLC17A4 A318T (rs11754288), p = 9.58 × 10\textsuperscript{-5}, GCKR L446P (rs1260326), p = 2.23 × 10\textsuperscript{-4} and ALDH2 E504K (rs671), p = 6.80 × 10\textsuperscript{-5}). We noticed the rarity of SLC2A9 V253I (rs16890799) and ABCG2 Q126X (rs72552713) (with a minor allele frequency of about 1%) in our sample. As the
All the OR (95% CI) and p values reported in this study were based on the PCA adjustment analysis. The variants with p < 0.05 were indicated in bold. All the OR (95% CI) and p values reported in this study were based on the PCA adjustment analysis.

Table 1. Replication of previously reported gout/SU GWAS associations in a cohort of Han Chinese. *1, insert; CHR, Chromosome; SNP, dbSNP rs number; BP, Position, based on hg19; A1, minor allele for the whole sample; Freq., frequency of A1 for cases/controls; OR, odds ratio, for A1; L95, the lower endpoint of the 95% confidence interval (CI); U95, the upper endpoint of the 95% confidence interval; P, p value. Reported gene(s), The reported gene(s) in the previous GWAS; aa_change, amino acid change; Gout or SU, indicating whether the locus found to be associated with gout or SU. The variants with p < 0.05 were indicated in bold. All the OR (95% CI) and p values reported in this study were based on the PCA adjustment analysis.

| CHR | SNP | BP    | A1       | Freq. | OR   | L95   | U95   | P          | Reported gene(s) | Gout or SU |
|-----|-----|-------|----------|-------|------|-------|-------|------------|------------------|------------|
| 2   | rs6547962 | 27,734,972 | A | 0.376/0.455 | 0.696 | 0.574 | 0.844 | 2.20E-04 | GCKR            | Gout, SU    |
| 4   | rs1722228 | 9,915,741  | T | 0.373/0.298 | 1.619 | 1.325 | 1.979 | 2.40E-06 | SLC2A9          | Gout, SU    |
| 4   | rs12505410 | 89,030,841  | G | 0.187/0.302 | 0.571 | 0.454 | 0.716 | 1.33E-06 | ABCG2           | Gout, SU    |
| 4   | rs1481012 | 89,039,082  | G | 0.498/0.305 | 1.890 | 1.559 | 2.291 | 8.96E-11 | ABCG2           | Gout, SU    |
| 6   | rs68094823 | 25,795,971  | P | 0.169/0.205 | 0.546 | 0.421 | 0.707 | 4.33E-06 | SLC17A1        | SU          |
| 11  | rs801733  | 65,934,549  | C | 0.015/0.021 | 0.428 | 0.202 | 0.905 | 0.0264 | CNH2            | Gout        |
| 12  | rs11066008 | 112,140,669 | G | 0.153/0.201 | 0.666 | 0.510 | 0.871 | 2.94E-03 | MYL2-CUX2 (ALDH2) | Gout |
| 17  | rs9895661 | 59,456,589  | C | 0.398/0.470 | 0.594 | 0.483 | 0.730 | 6.94E-07 | BCAS3           | Gout        |

Table 2. Association results for the selected important variants. CHR, Chromosome; SNP, dbSNP rs number; BP, Position, based on hg19; A1, minor allele for the whole sample; Freq., frequency of A1 for cases/controls; OR, odds ratio, for A1; L95, the lower endpoint of the 95% confidence interval (CI); U95, the upper endpoint of the 95% confidence interval; P, p value. Reported gene(s), The reported gene(s) in the previous GWAS; aa_change, amino acid change; Gout or SU, indicating whether the locus found to be associated with gout or SU. The variants with p < 0.05 were indicated in bold. All the OR (95% CI) and p values reported in this study were based on the PCA adjustment analysis.

| CHR | SNP | BP    | A1       | Freq. | OR   | L95   | U95   | P          | Reported gene(s) | Gout or SU |
|-----|-----|-------|----------|-------|------|-------|-------|------------|------------------|------------|
| 2   | rs1260326 | 27,730,940  | C | 0.378/0.456 | 0.698 | 0.576 | 0.845 | 2.23E-04 | GCKR (L446P)     | Gout, SU    |
| 2   | rs2307394  | 148,716,428 | T | 0.443/0.457 | 0.918 | 0.762 | 1.104 | 0.3628 | ORC4 (A785S)    | SU          |
| 4   | rs16890799 | 9,922,167    | T | 0.017/0.017 | 0.542 | 0.229 | 1.285 | 0.1644 | SLC2A9 (V235L)  | Gout, SU    |
| 4   | rs3775948  | 9,995,182    | G | 0.322/0.402 | 0.738 | 0.604 | 0.903 | 3.09E-03 | SLC2A9          | Gout, SU    |
| 4   | rs2231142  | 89,052,333   | T | 0.496/0.309 | 1.837 | 1.519 | 2.222 | 3.83E-10 | ABCG2 (Q141K)   | Gout, SU    |
| 4   | rs72552713 | 89,052,957   | A | 0.009/0.005 | 1.404 | 0.5316 | 3.707 | 0.4936 | ABCG2 (Q126X)   | Gout, SU    |
| 6   | rs11574288 | 25,776,949   | A | 0.155/0.184 | 0.582 | 0.443 | 0.764 | 9.58E-05 | SLC17A4 (A318T) | SU          |
| 6   | rs1165196  | 25,813,150   | G | 0.168/0.203 | 0.570 | 0.441 | 0.738 | 1.94E-05 | SLC17A1 (I269T) | SU          |
| 11  | rs4073582  | 66,050,712   | A | 0.013/0.020 | 0.431 | 0.198 | 0.938 | 0.0339 | CNH2            | Gout        |
| 12  | rs671     | 112,241,766  | A | 0.116/0.153 | 0.684 | 0.519 | 0.900 | 6.80E-03 | ALDH2 (E504K)   | Gout        |

minor allele frequency of ABCG2 Q126X is about 1%, the effect of ABCG2 Q141K will hide the effect by Q126X, we thus performed a multivariate logistic regression for Q126X and Q141K of ABCG2 (Supplementary Table S6). Comparing to the univariate analysis, the effect for Q126X was increased (the OR was increased from 1.404 to 2.027), which is consistent with result from similar analysis in the Japanese study. However, it remained non-significant (p = 0.1612). These rare variants often required a larger sample size for detecting significant associations. Similarly, SLC22A12 G65W (rs12800450) and ALDH16A1 G65W (rs150414818) were absent in our dataset. Both were low-frequency variants identified to be associated with gout in European and/or Americans samples, however, they were non-polymorphic in the 1000 Genomes Project datasets. Of noted, the other gout associated SNPs identified in the Japanese study also showed direction-consistent association and with nominal significance in our dataset (rs4073582, p = 0.0339 and rs3775948, p = 3.09 × 10⁻³).

We further investigated the cumulative effect for risk alleles of gout associated variants at these loci. Conditional analysis was used to test independent effect for the loci with multiple significant SNPs. The previously identified SNPs (especially the gout associated and non-synonymous ones) were given higher priorities in the analysis for their more important roles. The conditional analysis indicated seven independent variants for the gout associated loci: rs1260326 (L446P) of GCKR, rs11722228 of SLC2A9, rs12505410 and rs2231142 (Q141K) of ABCG2, rs4073582 of CNH2, rs671 (E504K) of ALDH2 (MYL2-CUX2) and rs9895661 of BCAS3 (Supplementary Table S7), thus we only included these independent variants in the cumulative genetic risk score analysis. We observed a strong increase in the OR with increasing risk allele load (Fig. 1). Comparing to the reference category of having five or fewer risk alleles, ORs for having 8, 9, 10, 11 or 12 more risk alleles were 1.310, 2.925, 4.158, 6.892 and 16.361, respectively (Supplementary Methods and Table S8).

One of the 12 significant variants from the SU associated loci, rs68094823 (p = 4.33 × 10⁻⁶, OR = 0.546), was statistically significant after Bonferroni correcting (Table 1). Rs68094823 is an intron variant of SLC17A1 (also known as NPT1), and it’s in strong LD with a previously identified SU associated variant (rs1165151,
with renal underexcretion gout (a major subtype of gout), but not significant for all gout\(^2\). In the present study, rs1165196 (I269T), required special attention. A previous study showed rs1165196 was significantly associated with SU-raising allele. For the SLC17A1 locus, a common missense variant,\(^{14,22}\) the gout risk allele is in highly LD with the SU-raising allele. For the MYL2-CUX2-GCKR (ABCG2) loci\(^14,22\) associated with gout\(^2,9,11,14,25\). We, for the first time, replicated the associations of the well-known urate transporter genes (\(SLC2A9\) and \(ABCG2\)) were included and \(r^2 = 0.90\)^11. Haplotype analysis suggested that our finding were consistent with previous finding, that is, the gout risk allele is in highly LD with the SU-raising allele. For the \(SLC17A1\) locus, a common missense variant, rs1165196 (I269T), required special attention. A previous study showed rs1165196 was significantly associated with renal underexcretion gout (a major subtype of gout), but not significant for all gout\(^2\). In the present study, we provided statistically significant evidence for rs1165196 (\(p = 1.94 \times 10^{-5}\), OR = 0.570), thus we confirmed the association of rs1165196 with gout (Table 2). The conditional analysis showed that rs1165196 could be the one independent variant in the \(SLC17A1\) locus (Supplementary Table S7).

For the SU associated loci, we observed 12 independent variants (rs17632159, rs6935612, rs1165196 (\(SLC17A1\) I269T), rs3734692, rs9321446, rs9314273, rs10821871, rs2361216, rs11172134, rs7978353, rs61168554 and rs11150190). In the cumulative genetic risk score analysis of these variants, we also observed a trend of increase in risk for gout with the growing number of the risk alleles (Supplementary Methods and Figure S1). When setting the reference group as having eight or fewer risk alleles, ORs for the groups having more risk alleles ranged from to 1.644 to 8.884 (Supplementary Table S9). Additionally, we also found an additive effect of the variants from the gout and SU associated loci. The tendency of increasing ORs for cumulative effect of seven variants on gout associated loci escalated, when additional risk alleles on SU associated loci were considered (Fig. 1). Comparing to the reference category as having five or fewer risk alleles at the variants on gout associated loci, ORs ranged from 1.697 to 30.230 for the categories having 8 or more risk alleles on gout associated loci, and at the same time having nine or more risk alleles on SU associated loci (Supplementary Methods and Table S10).

**Discussion**

We used a Han Chinese GWAS data of clinically defined gout cases to investigate whether variants associated with gout/SU in other studies can be replicated. For the previously reported gout associated loci, rs1260326 (L446P)of GCKR, rs1722228 of SLC2A9, rs12505410 and rs2231142 (Q141K)of ABCG2, rs4073582 of CNIH2, rs671 (E504K) of ALDH2 (MYL2-CUX2) and rs9895661 of BCAS3 were included and eight bins (≤5, 6, 7, 8, 9, 10, 11, and ≥12) were generated. Using the ≤5 bin as the reference category, the OR and 95% CI for each of the other bins (6, 7, 8, 9, 10, 11, and ≥12) were assessed using logistic regression. For the combined analysis of variants from gout and SU associated loci (GOUT + SU, red color), we also used the ≤5 bin in the gout associated loci analysis as reference, and excluded the individuals with ≤8 risk alleles in the SU associated loci analysis from the test bins (Supplementary Methods).

**Figure 1.** Cumulative effect of the associated variants from gout associated loci and gout + SU associated loci on gout incidence. For the analysis using variants from the gout associated loci (GOUT, blue color), seven variants (rs1260326 (L446P)of GCKR, rs1722228 of SLC2A9, rs12505410 and rs2231142 (Q141K)of ABCG2, rs4073582 of CNIH2, rs671 (E504K) of ALDH2 (MYL2-CUX2) and rs9895661 of BCAS3) were included and eight bins (≤5, 6, 7, 8, 9, 10, 11, and ≥12) were generated. Using the ≤5 bin as the reference category, the OR and 95% CI for each of the other bins (6, 7, 8, 9, 10, 11, and ≥12) were assessed using logistic regression. For the combined analysis of variants from gout and SU associated loci (GOUT + SU, red color), we also used the ≤5 bin in the gout associated loci analysis as reference, and excluded the individuals with ≤8 risk alleles in the SU associated loci analysis from the test bins (Supplementary Methods).
References

1. Whitfield, J. B. & Martin, N. G. Inheritance and alcohol as factors influencing plasma uric acid levels. *Acta geneticae medicae et genetologiae* 32, 117–126 (1983).

2. Li, S. et al. The GLUT9 gene is associated with serum uric acid levels in Sardinia and Chianti cohorts. *PLoS genetics* 3, e194, doi:10.1371/journal.pgen.0030194 (2007).

3. Dehghan, A. et al. Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study. *Lancet* 372, 1953–1961, doi:10.1016/S0140-6736(08)61343-4 (2008).

4. Doring, A. et al. SLC2A9 influences uric acid concentrations with pronounced sex-specific effects. *Nature genetics* 40, 430–436, doi:10.1038/ng.107 (2008).

5. Vitar, V. et al. SLC2A9 is a newly identified urate transporter influencing serum urate concentration, urate excretion and gout. *Nature genetics* 40, 437–442, doi:10.1038/ng.106 (2008).

6. Wallace, C. et al. Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *American journal of human genetics* 82, 139–149, doi:10.1016/j.ajhg.2007.11.001 (2008).

7. Kamatani, Y. et al. Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nature genetics* 42, 210–215, doi:10.1038/ng.531 (2010).

8. Yang, Q. et al. Multiple genetic loci influence serum urate levels and their relationship with gout and cardiovascular disease risk factors. *Circulation. Cardiovascular genetics* 3, 523–530, doi:10.1161/CIRCGENETICS.109.934555 (2010).

9. Sulem, P. et al. Identification of low-frequency variants associated with gout and serum uric acid levels. *Nature genetics* 43, 1127–1130, doi:10.1038/ng.972 (2011).

10. Tin, A. et al. Genome-wide association study for serum urate concentrations and gout among African Americans identifies genomic risk loci and a novel URAT1 loss-of-function allele. *Human molecular genetics* 20, 4056–4068, doi:10.1093/hmg/ddr307 (2011).

11. Kottingen, A. et al. Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nature genetics* 45, 143–154, doi:10.1038/ng.2500 (2013).

12. Yang, B. et al. A genome-wide association study identifies common variants influencing serum uric acid concentrations in a Chinese population. *BMC medical genomics* 7, 10, doi:10.1186/1755-8794-7-10 (2014).

13. Li, C. et al. Genome-wide association analysis identifies three new risk loci for gout arthritis in Han Chinese. *Nature communications* 6, 7041, doi:10.1038/ncomms8041 (2015).

14. Matsuo, H. et al. Genome-wide association study of clinically defined gout identifies multiple risk loci and its association with clinical subtypes. *Annals of the rheumatic diseases* 75, 652–659, doi:10.1136/annrheumdis-2014-206191 (2016).

15. Kolz, M. et al. Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. *PLoS genetics* 5, e1000504, doi:10.1371/journal.pgen.1000504 (2009).

16. Tu, H. P. et al. Associations of a non-synonymous variant in SLC2A9 with gouty arthritis and uric acid levels in Han Chinese subjects and Solomon Islanders. *Annals of the rheumatic diseases* 69, 887–891, doi:10.1136/ard.2009.113357 (2010).

17. Li, C. et al. Multiple single nucleotide polymorphisms in the human urate transporter 1 (hURAT1) gene are associated with hyperuricemia in Han Chinese. *Journal of medical genetics* 47, 204–210, doi:10.1136/jmg.2009.068619 (2010).

18. Wang, J. et al. Association between gout and polymorphisms in GCKR in male Han Chinese. *Human genetics* 131, 1261–1265, doi:10.1007/s00439-012-1151-9 (2012).

19. Wang, B. et al. Genetic analysis of ABCG2 gene C421A polymorphism with gout disease in Chinese Han male population. *Human genetics* 127, 245–246, doi:10.1007/s00439-009-0760-4 (2010).

20. Chiba, T. et al. NPT1/SLC17A1 Is a Renal Urate Exporter in Humans and Its Common Gain-of-Function Variant Decreases the Risk of Renal Underexcretion Gout. *Arthritis & Rheumatology* 67, 281–287, doi:10.1002/art.38884 (2015).

21. Matsuo, H. et al. Common Defects of ABCG2, a High-Capacity Urate Exporter, Cause Gout: A Function-Based Genetic Analysis in a Japanese Population. *Journal of medical genetics* 81, 559–575 (2007).

22. Sakiyama, M. et al. Identification of rs671, a common variant of ALDH2, as a gout susceptibility locus. *Scientific Reports* 6, doi:10.1038/srep25360 (2016).

23. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics* 81, 559–575 (2007).

24. World Medical, A. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *Jama* 310, 2191–2194, doi:10.1001/jama.2013.281053 (2013).

25. Woodward, O. M. et al. Identification of a urate transporter, ABCG2, with a common functional polymorphism causing gout. *Proceedings of the National Academy of Sciences of the United States of America* 106, 10338–10342, doi:10.1073/pnas.0901249106 (2009).

Acknowledgements

We would like to thank all the participants involved in this study. We also thank editors and anonymous reviewers for their valuable comments on the manuscript. This work was supported by the Ministry of Science and Technology of China (2016YFC0904000), the National Science Foundation of China (81520108007, 31371372, 81500346, 31471195, 81100621, 31325014, 81130022, 81272302 and 81421061). Science and Technology Development Project of Shandong Province (2014GSF118013), the 973 Program (2010CB534902, 2015CB559100), the 863 project (2012AA02A515), the China Postdoctoral Science Foundation (2016MS509615), the Shandong Postdoctoral Innovation Foundation (201601015), the Qingdao Postdoctoral Application Research Project (2016048), the Shanghai Jiao Tong Univ Liberal Arts and Sciences Cross-Disciplinary Project (13JCRZ02), “Shu Guang” project supported by Shanghai Municipal Education Commission, Shanghai Education Development Foundation (12SG17).

Author Contributions

Li C.G., Shi Y.Y. conceived, designed and led the study. Shi Y.Y., Li C.G., Li Z.Q., Zhou Z.W., Hou X. and Lu D.J. interpreted the main findings. Shi Y.Y., Li C.G., Li Z.Q. and Zhou Z.W. drafted the manuscript. Li Z.Q., Zhou Z.W., Hou X., Lu D.J., Yuan X., Han L. and Cui L.L. processed the bioinformatics/statistical analysis. Lu D.J., Lu J., Hou X., Chen J.H., Ji J., Wang C., Cheng X.Y., Zhang K.K. and Zhou Z.W. undertook the main experiments and raw data management. Li C.G., Hou X., Liu Z., Jia Z.T., Ma L.D., Xin Y., Liu T., Yu Q., Ren W., Wang X.F., Mi Q.S. and Li X.D. were responsible for the sample collection and clinical data management. Li C.G. and Shi Y.Y. obtained the main funding supports.
Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-04127-4

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017