Genetic variants in two pathways influence serum urate levels and gout risk: a systematic pathway analysis

Zheng Dong1, Jingru Zhou1, Xia Xu2, Shuai Jiang1, Yuan Li3, Dongbao Zhao2, Chengde Yang3, Yanyun Ma4, Yi Wang4, Hongjun He5, Hengdong Ji6, Juan Zhang7, Ziyu Yuan7, Yajun Yang1,7, Xiaofeng Wang1,7, Yafei Pang2, Li Jin1,7, Hejian Zou6,9 & Jiucun Wang1,7,9

The aims of this study were to identify candidate pathways associated with serum urate and to explore the genetic effect of those pathways on the risk of gout. Pathway analysis of the loci identified in genome-wide association studies (GWASs) showed that the ion transmembrane transporter activity pathway (GO: 0015075) and the secondary active transmembrane transporter activity pathway (GO: 0015291) were both associated with serum urate concentrations, with \( P_{\text{FDR}} \) values of 0.004 and 0.007, respectively. In a Chinese population of 4,332 individuals, the two pathways were also found to be associated with serum urate (\( P_{\text{FDR}} = 1.88 \times 10^{-5} \) and \( 3.44 \times 10^{-4} \), separately). In addition, these two pathways were further associated with the pathogenesis of gout (\( P_{\text{FDR}} = 1.08 \times 10^{-8} \) and \( 2.66 \times 10^{-3} \), respectively) in the Chinese population and a novel gout-associated gene, SLC17A2, was identified (OR = 0.83, \( P_{\text{FDR}} = 0.017 \)). The mRNA expression of candidate genes also showed significant differences among different groups at pathway level. The present study identified two transmembrane transporter activity pathways (GO: 0015075 and GO: 0015291) were associations with serum urate concentrations and the risk of gout. SLC17A2 was identified as a novel gene that influenced the risk of gout.

Uric acid is the final catabolic product of purine oxidation in humans. Because of the loss of important genes in the urate degradation pathway during human evolution, uric acid cannot be metabolized into allantoin1,2. It is primarily produced in the liver, excreted via the kidney and the gut, and present in the blood as urate. Elevated serum urate (hyperuricemia) is caused by an imbalance between uric acid production and disposal3. This imbalance results in the deposition of monosodium urate (MSU) crystals in the joints leading to gout. Gout is one of the forms of common inflammatory arthritis, and it affects nearly 4% of American adults4. Recent epidemiological studies have suggested that the prevalence and incidence of gout are increasing4–6. In addition, gout is often associated with other diseases, such as hypertension, obesity, renal insufficiency, and heart failure7.

Serum urate concentrations exhibit a strong genetic predisposition, with a heritability estimate ranging from 40% to 70%8. Genome-wide association studies (GWASs) have identified several genetic variant loci in various genes associated with serum urate concentrations8–15. However, those genetic variants only explain 7.0% of the variance in serum urate concentrations8, which is distinctly less than what is estimated to be heritable. The key challenges for these GWASs are identifying causal SNPs and providing abundant evidence for the influence of candidate loci on serum urate concentrations by specific biological mechanisms16. Moreover, multiple genes may

---

1State Key Laboratory of Genetic Engineering, Collaborative Innovation Center for Genetics and Development, School of Life Sciences, Fudan University Jiangwan Campus, Shanghai, China. 2Division of Rheumatology and Immunology, Changhi Hospital, Shanghai, China. 3Division of Rheumatology, Ruijin Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China. 4Ministry of Education Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai, China. 5Division of Rheumatology, Taixing People’s Hospital, Jiangsu Province, China. 6Division of Rheumatology, Taizhou People’s Hospital, Jiangsu Province, China. 7Fudan-Taizhou Institute of Health Sciences, Taizhou, Jiangsu Province, China. 8Division of Rheumatology, Huashan Hospital, Fudan University, Shanghai, China. 9Institute of Rheumatology, Immunology and Allergy, Fudan University, Shanghai, China. Zheng Dong, Jingru Zhou, Xia Xu and Shuai Jiang contributed equally to this work. Correspondence and requests for materials should be addressed to J.W. (email: jcwang@fudan.edu.cn)
interact with each other and may work together to influence the development of a specific disease, especially in the case of complex disorders. In addition, pathway analysis has been shown to be a useful approach for clarifying biological insights and identifying new candidate genes. Therefore, it is necessary to identify additional genetic factors influencing serum urate concentrations and the pathogenesis of gout using pathway analysis.

In the present study, we aimed to identify serum urate- and gout-associated candidate genes and biological pathways using a four-step approach. First, we applied pathway analysis to the SNPs identified as associated with serum urate concentrations in previous GWASs to identify candidate pathways and candidate genes in pathways. Then, the candidate genes and pathways were tested in 4,332 Chinese individuals to validate the associations between candidate genes and pathways and serum urate concentrations. Next, those candidate genes and pathways were also tested in a case-control study to find the candidate genes and pathways affecting the risk of gout. Finally, the transcription levels of candidate genes were tested at the pathway level. Using this strategy, this study identified two transmembrane transporter activity-related pathways that influenced serum urate levels and the pathogenesis of gout.

| Pathway Function | Candidate genes | \( P \) | \( P_{\text{FDR}} \) |
|------------------|-----------------|--------|---------------|
| GO:0015075 ion transmembrane transporter activity | PKD2, SLC2A9, SLC17A3, SLC22A11, SLC17A1, SLC17A4, SLC5A6, SLC17A2 | <0.001 | 0.004 |
| GO:0015291 secondary active transmembrane transporter activity | SLC2A9, SLC17A3, SLC22A11, SLC17A1, SLC17A4, SLC5A6, SLC16A9, SLC17A2 | <0.001 | 0.007 |

Table 1. Candidate causal pathways of serum urate identified by GWAS data. \( P_{\text{FDR}} \) means \( P \) value adjusted for multiple comparisons correction using FDR method.

Results

Pathway analysis for serum urate in GWAS datasets. The pathway analysis of the GWAS data from European and US populations identified nine candidate causal genes (PKD2, SLC2A9, SLC17A3, SLC22A11, SLC17A1, SLC17A4, SLC5A6, SLC17A2 and SLC16A9) and two candidate causal pathways (the ion transmembrane transporter activity pathway (GO: 0015075) and the secondary active transmembrane transporter activity pathway (GO: 0015291)) associated with serum urate concentration (\( P_{\text{FDR}} = 0.004 \) and 0.007) (Table 1). Among the nine candidate genes, SLC17A2 and SLC5A6 were two novel genes associated with serum urate concentration (Table 1 and Table S1). The nine candidate genes were distinctly clustered in the two candidate causal pathways, and the genes of the two pathways revealed some overlap. For example, SLC2A9, SLC17A3, SLC22A11, SLC17A1, SLC17A4, SLC5A6 and SLC17A2 were involved in both the two pathways, while PKD2 and SLC16A9 were only involved in one pathway each (Table 1). Additional information about the differences and the commonalities of the genes in the two candidate causal pathways are shown in Table S2.

Candidate causal pathways associated with serum urate and gout. To verify the association between the two candidate pathways and the serum urate concentration in the Chinese population and to further analyze their effects on the pathogenesis of gout, approximately 5,000 Chinese individuals were enrolled in this study. Four candidate genes (PKD2, SLC2A9, SLC17A3 and SLC17A1) were verified to be associated with serum urate with \( P_{\text{FDR}} \) values less than 0.05 (Table 2). However, SLC17A2, SLC16A9, SLC22A11, SLC17A4 and SLC5A6 were not associated with serum urate concentration in the Chinese population, suggesting that ethnicity might be a heterogeneity factor for the associations. Both the two candidate pathways were confirmed to affect the concentration of serum urate (GO: 0015075, \( P_{\text{FDR}} = 1.88 \times 10^{-5} \); GO: 0015291, \( P_{\text{FDR}} = 3.44 \times 10^{-4} \), respectively).

In addition, five genes (PKD2, SLC2A9, SLC17A1, SLC17A4 and SLC17A2) were determined to influence the risk of gout (all \( P_{\text{FDR}} < 0.05 \) (Table 2). To avoid the heterogeneity of gender, logistic regression adjusted for gender was performed and the result also showed that those five genes exhibited effects on the susceptibility of gout (all \( P_{\text{FDR}} < 0.05 \)). SLC17A2 was a novel gene that was found to be associated with the pathogenesis of gout (OR = 0.83, \( P_{\text{FDR}} = 0.017 \)). SLC17A4 was also identified as a candidate gene for gout risk (OR = 0.71, \( P_{\text{FDR}} = 5.16 \times 10^{-4} \)), which was consistent with the result of our previous study. In addition, two serum urate-associated pathways were found to affect the risk of gout (GO: 0015075, \( P_{\text{FDR}} = 1.88 \times 10^{-5} \); GO: 0015291, \( P_{\text{FDR}} = 2.66 \times 10^{-3} \), respectively), indicating a key role for elevated uric acid in the development of gout. To the best of our knowledge, this is the first time that these two candidate pathways have been associated with gout risk. To further understand the pathogenetic mechanism linking serum urate and gout, we studied the genetic effects in transition process from hyperuricemia to gout. The result showed that PKD2, SLC17A1, SLC17A4 and SLC17A2 were associated with this pathogenesis (all \( P_{\text{FDR}} < 0.05 \) (Table 2), while the two candidate pathways had no effect on pathogenesis in the pathway analysis (GO: 0015075, \( P_{\text{FDR}} = 0.158 \); GO: 0015291, \( P_{\text{FDR}} = 0.111 \), respectively).

As expected, gender was a heterogeneity factor for serum urate and gout that affected the associations identified in the present study. Distinctly different results for those associations were observed between males and females (Table 2). For example, the two candidate pathways influenced the concentration of serum urate and gout risk in males (serum urate: GO: 0015075, \( P_{\text{FDR}} = 3.84 \times 10^{-4} \); GO: 0015291, \( P_{\text{FDR}} = 5.24 \times 10^{-4} \), gout: GO: 0015075, \( P_{\text{FDR}} = 1.22 \times 10^{-7} \); GO: 0015291, \( P_{\text{FDR}} = 1.83 \times 10^{-3} \), respectively), but not in females (serum urate: GO: 0015075, \( P_{\text{FDR}} = 0.538 \); GO: 0015291, \( P_{\text{FDR}} = 0.317 \), gout: GO: 0015075, \( P_{\text{FDR}} = 1.00 \); GO: 0015291, \( P_{\text{FDR}} = 1.00 \), respectively).

Associations between candidate pathways and serum urate in the subgroups of BMI. Obesity is another heterogeneity factors that have been shown to influence serum urate concentration, however, the effects of obesity on the associations between the candidate pathways and urate concentration were limited. Therefore, further analyses of subgroups of individuals divided by BMI was performed.
Table 2. Association between genetic variants in pathways and serum urate, hyperuricemia and gout. MAF, minor allele frequency in controls. HUA, hyperuricemia. $\beta$ values for SNP in serum urate were calculated by linear regression adjusted age and gender. $P$ values for SNP in serum urate were calculated by deviance analysis for linear regression adjusted age and gender. $P$ values for SNP in hyperuricemia and gout were calculated by Fisher’s exact test. $P_{\text{FDR}}$ value for SNPs and pathways were multiple corrected by FDR method. $P$ values for pathway were calculated by Simes test.

Based on the analysis of the BMI subgroups, four genes (PKD2, SLC17A3, SLC17A1 and SLC17A4) were found to be associated with serum urate concentration in individuals with normal weight (18.5 $\leq$ BMI < 25) with a $P_{\text{FDR}}$ less than 0.05 and two genes (PKD2 and SLC2A9) were found to be associated with serum urate concentration in overweight subjects (BMI $\geq$ 25) (Table 3). In addition, the ion transmembrane transporter activity pathway (GO: 0015075) affected the concentrations of serum urate in both normal and overweight people ($P_{\text{FDR}} = 0.018$ and $P_{\text{FDR}} = 2.15E-03$, respectively), while the secondary active transmembrane transporter activity pathway (GO: 0015291) only affected the concentration of serum urate in overweight individuals with a $P_{\text{FDR}}$ equal to 2.15E-03.

The contribution of genetic effects on the pathogenesis of gout. Because all candidate loci and pathways only influenced the concentrations of serum urate and the risk of gout in males (Table 2), further analysis for the candidate loci and pathways only preformed in males. Across the nine candidate loci for nine genes in the two pathways (rs2728121, rs13129697, rs2242206, rs1165165, rs3759053, rs1165196, rs11754288, rs1395 and rs2071299), for each additional effect allele in males, the genetic effect on serum urate was positively linearly correlated with the odds ratio for gout ($R^2 = 0.772$) (Fig. 1A), which was consistent with the fact that high serum urate concentration is a key risk factor for the development of gout17. To measure the contribution of the
development from hyperuricemia to gout, we analyzed the correlation between the genetic effects on the associations of gout and hyperuricemia/control. The result showed a high and linear correlation between those two combinations ($R^2 = 0.803$), indicating the importance of this development in the pathogenesis of gout (Fig. 1B).

**Weight and unweighted genetic urate risk score of candidate pathways associated with hyperuricemia and gout.** The weighted and unweighted genetic urate score (GRS) for candidate loci in the ion transmembrane transporter activity pathway (GO: 0015075) explained an average of 1.47% and 0.96% of the variance in serum urate concentration and was strongly associated with hyperuricemia and gout in males (weighted GRS: coefficients $= 0.014$, $P = 2.53E-07$; coefficients $= 0.025$, $P = 2.14E-07$, respectively; unweighted GRS: coefficients $= 0.081$, $P = 4.27E-05$; coefficients $= 0.172$, $P = 1.49E-06$, respectively). The weighted and unweighted GRS for the secondary active transmembrane transporter activity pathway (GO: 0015291) explained 0.81% and 0.45% of the variance in serum urate concentration and was found to affect hyperuricemia and gout in males (weighted GRS: coefficients $= 0.013$, $P = 1.21E-04$; coefficients $= 0.016$, $P = 4.09E-03$, respectively; unweighted GRS: coefficients $= 0.059$, $P = 0.003$; coefficients $= 0.085$, $P = 0.013$, respectively). Furthermore, the increased genetic urate score for the two candidate pathways resulted in elevated hyperuricemia and gout in males (Fig. 2). The above results suggested that the two candidate pathways influenced the pathogenesis of gout by modifying the level of serum urate.

**The different expression levels of candidate genes among the groups.** To further validate the results described above, we analyzed the differences in transcription levels of the candidate genes in 70 male gout patients, 85 male hyperuricemia patients and 58 male healthy individuals (Fig. 3). Six candidate genes (SLC17A1, SLC17A4 and SLC2A9) were not under consideration in this section of result due to their mRNA expressions being too high (Figs 4, S1 and S2). The result showed the SLC17A2 protein level was higher in hyperuricemia patients than in gout patients and healthy control groups.

**Discussion**

In the present study, we analyzed the genetic variants in candidate pathways affecting the concentration of serum urate and the risk of gout, and we identified two candidate pathways. All candidate loci and pathways involved in the present studies were analyzed in detail, including the analysis of associations in subgroups based on gender and body mass index. For each effect allele of the candidate genes, the genetic effect on uric acid was positively linearly correlated with the gout odds ratio. The expression of candidate genes was also analyzed at pathway level.

Nine candidate causal genes (PKD2, SLC2A9, SLC17A3, SLC22A11, SLC17A1, SLC17A4, SLC5A6, SLC17A2 and SLC16A9) and two candidate causal pathways (the ion transmembrane transporter activity pathway (GO: 0015075) and the secondary active transmembrane transporter activity pathway (GO: 0015291)) were found to be associated with serum urate concentration in the pathway analysis of the GWAS dataset. Among these genes, SLC17A2 and SLC5A6 were two novel genes found to be associated with serum urate concentration in European

| Subgroup | Variants | Gene | Effect allele | Underweight | Normal | Overweight |
|----------|----------|------|---------------|-------------|--------|------------|
| BMI      | rs2728121 | PKD2| C             | $-13.03$    | 0.636  | 0.954      |
|          | rs13129697 | SLC2A9| T            | 21.05       | 0.235  | 0.424      |
|          | rs22242206 | SLC16A9| G             | $-17.47$    | 0.202  | 0.455      |
|          | rs165165 | SLC17A3| T            | 24.86       | 0.062  | 0.561      |
|          | rs3759053 | SLC22A11| T           | $-9.15$     | 0.666  | 0.856      |
|          | rs165196 | SLC17A1| G            | $-24.38$    | 0.096  | 0.287      |
|          | rs1754288 | SLC17A4| A            | $-28.94$    | 0.089  | 0.400      |
|          | rs1395  | SLC5A6| G             | 1.26        | 0.990  | 0.990      |
|          | rs2071299 | SLC17A2| G            | $-1.79$     | 0.865  | 0.973      |

Table 3. Association between genetic variants in pathways and serum urate in subgroups of BMI. Subgroup of BMI: 1, Underweight (BMI < 18.5); 2, Normal (18.5 ≤ BMI < 25); 3, Overweight (BMI ≥ 25). $P$ values for SNP in serum urate were calculated by linear regression adjusted age and gender. $P$ values for SNP in serum urate were calculated by deviation analysis for linear regression adjusted age and gender. $P$ values for pathway were calculated by Simes test. $P_{FDR}$ values for SNPs and pathways were multiple corrected by FDR method.
and US individuals; however, this association was not replicated in the Chinese population, illustrating the heterogeneity effect of ethnicity on those associations. However, the transcription levels of both SLC17A2 and SLC5A6 were significantly different between the hyperuricemia patients and healthy subjects, suggesting that these two genes potentially affect serum urate levels.

The protein coded by SLC5A6, sodium dependent multivitamin transporter (SMVT), is a transmembrane protein that mediates the cellular uptake of biotin, lipoic acid and pantothenic acid. Biotin is an important cofactor for several carboxylase enzymes that are involved in the metabolism of carbohydrates, amino acids, and fatty acids. Pantothenic acid is an essential component for the synthesis of coenzyme A, which is required for cellular metabolism and plays an important role in the tricarboxylic acid cycle (TCA cycle). Tricarboxylic acid cycle intermediates serve as substrates for purine synthesis, and purine metabolism can directly influence the concentration of serum urate. Taken together, SLC5A6 might affect the concentrations of serum urate by mediating the cellular uptake of biotin, lipoic acid and pantothenic acid.

SLC17A2 (NPT3) belongs to a functional subgroup of the SLC17 family, type I phosphate transporters, which also includes SLC17A1, SLC17A3 and SLC17A4. SLC17A1-4 are only found in mammals and originated from a common ancestor, suggesting a similar urate transporter function of these four genes. A recent study in mouse showed that NPT3 transports urate was similar with other NPTs. In addition, SLC17A2 was expressed in the liver and the kidney, which are important organs in the regulation of serum urate concentration. Urate is synthesized predominantly in the liver, and around two-thirds of urate excretion occurs via the kidneys. In the present study, SLC17A2 was found to be associated with gout risk either in the presence of hyperuricemia or not, and it variant played a protective role against gout. The common variant rs2071299 is an expression associated SNP and SLC17A2 mRNA expression is modified by different genotypes of this variant. Based on the prediction of functional effects of this variant by F-SNP, rs2071299 can cause frameshift coding and splice site change. In our study, the mRNA expression and protein levels of SLC17A2 between healthy individuals and hyperuricemia patients were significantly different, also suggesting potential correlation between this variant and mRNA expression. Thus, SLC17A2 might influence the development of gout by changing its expression.

Figure 1. Positive correlation between the genetic effect on serum urate concentration and gout in males. (A) urate beta coefficients and gout odds ratios; (B) odds ratios for gout vs. hyperuricemia and gout odds ratios. Each confidence interval for the beta coefficient or odds ratio estimate was plotted as a bar of the point.
Both the two candidate pathways identified in the present pathway analysis of GWAS data from the European and US population were found to affect the concentration of serum urate in the Chinese population. The function of the two candidate pathways involved transmembrane transporter activity, and several genes in these pathways had been identified to be urate transporters, such as \( SLC2A9 \), \( SLC17A1 \) and \( SLC17A4 \). However, another famous urate transporter, \( ABCG2 \), was not found in the two identified pathways, indicating that additional pathways that function in urate transport should be identified in future studies. Because elevated serum urate is a key
risk factor for gout, we analyzed the associations between the two urate-associated pathways and gout, and we identified significant associations. In addition, the mRNA expression of candidate genes analyzed at pathway level also showed significant differences. To the best of our knowledge, this is the first time these two pathways have been associated with the pathogenesis of gout.

Heterogeneity factors, such as gender and BMI, can influence the associations of genetic variants and serum urate/gout, as reported in our previous study. In the present study, the two candidate pathways played important roles in regulating the concentration of serum urate and in the development of gout in males, but not in females, which was consistent with the different effects of genetic variants between males and females. Regarding BMI, the ion transmembrane transporter activity pathway affected the concentration of serum urate in both normal and overweight individuals, while the secondary active transmembrane transporter activity pathway only affected the concentration of serum urate in overweight subjects, suggesting that those associations were strongly influenced by BMI.

There are several limitations that should be considered in the present study. First, the number of female gout patients was limited. Second, the two candidate pathways identified above could only explain a portion of the variance of serum urate concentrations, and more urate-associated pathways should be identified in future studies. Finally, although gender and BMI were assessed in this study, some other heterogeneity factors associated with urate and gout were not considered. Therefore, further studies using pathway analysis should be performed.

Taken together, the present study identified the ion transmembrane transporter activity pathway (GO: 0015075) and the secondary active transmembrane transporter activity pathway (GO: 0015291) as associated with serum urate concentrations and the pathogenesis of gout. To the best of our knowledge, this is the first time these two pathways have been associated with gout. In addition, we identified a novel gout-associated gene, SLC17A2, in this study. These findings contributed to our understanding of the biological mechanisms regulating serum urate and gout on a pathway level and demonstrated the importance of analyzing GWAS data by pathway analysis.

Materials and Methods

**Genome-wide association studies (GWASs) dataset.** Serum urate-associated SNPs identified in two meta-analysis of genome-wide association studies (GWASs) for European and US population were analyzed in the present study. One study, comprised of 110,347 participants from 48 studies of serum urate concentrations, found that the median serum urate concentration was 5.2 mg/dl (the range was from 3.9 to 6.1 mg/dl). Another GWAS meta-analysis was performed in 5 population-based cohorts, which included 28,283 participants, for serum urate concentrations and found that the median (SD) of serum urate concentrations in the 5 cohorts were 354.8 \(\mu\)mol/l (96.3), 353.9 \(\mu\)mol/l (89.4), 327.4 \(\mu\)mol/l (85.7), 315.2 \(\mu\)mol/l (89.22), 321.2 \(\mu\)mol/l (80.9) and 324.3 \(\mu\)mol/l (117.2), respectively. The descriptions of the GWASs included in this meta-analysis have been described in detail elsewhere. The two studies described above were approved by the corresponding local ethics committees and written informed consent was obtained from the study participants.

**Pathway analysis for GWAS data.** We used ICSNPathway to analyze GWAS datasets on serum urate concentrations. ICSNPathway offers a two-stage analysis for GWAS data. In the first stage, it pre-selects potential causal SNPs from the most significant SNPs in the GWAS dataset using linkage disequilibrium (LD) analysis and functional SNP annotation. In the second stage, it applies a pathway-based analysis (PBA) to identify the specific biological mechanisms for the pre-selected potential causal SNPs. Two GWAS datasets containing 1,279 and 2,201 SNPs associated with serum urate concentrations \((P \leq 4.00E-07 \text{ and } P < 5.00E-08, \text{ respectively})\) were used as the inputs for the ICSNPathway analysis. The main parameters used in this study included: (a) threshold to extract the most significant SNPs: \(P\) value < \(10^{-6}\); (b) HapMap population: European American (CEU) (HapMap 3); (c) LD cutoff: \(r^2 > 0.8\); (d) the maximum distance to search LD neighborhoods: 200 kb; (e) the method used for mapping a SNP to its corresponding gene (s): 20 kb upstream and downstream of the gene; (f) pathway set database: KEGG (http://www.genome.jp/kegg/pathway.html), BioCarta (http://www.biocarta.com/genes/index.asp), GO biological process (gene ontology, http://www.geneontology.org) and GO molecular function (http://www.broadinstitute.org/gsea/msigdb/index.jsp); (g) number of genes in each pathway: minimum 5 and maximum 100; and (h) false discovery rate (FDR) cutoff for multiple testing correction for PBA: \(P_{\text{FDR}} < 0.05\) was considered statistically significant.

![Figure 4](image-url) The protein levels of SLC17A2 were determined by western blot and analyzed by densitometry. (A) Western blot analysis of SLC17A2 protein content in control, hyperuricemia, gout groups and (B) densitometry analysis of western blots for SLC17A2. GAPDH was used as an internal control. The full-length gels are presented in Supplementary Figures 1 and 2. Data are presented as mean ± SD of five samples and compared with Student’s t test; *\(P < 0.05\), **\(P < 0.01\).
Participants. This study was approved by the Ethical Committees of the School of Life Sciences of Fudan University (approval number of 140) and was conducted in accordance with the guidelines and regulations of the Declaration of Helsinki. All participants provided written informed consent. A total of 582 gout patients who were clinically diagnosed according to the American College of Rheumatology diagnostic criteria were enrolled from Changhai Hospital, Taizhou People's Hospital and Taixing People's Hospital. An additional 4,332 healthy individuals without a history of gout were enrolled from the Taizhou Longitudinal Study.

The 4,332 individuals were divided into subgroups of patients with hyperuricemia and healthy controls according to the threshold of serum urate deposition, i.e., approximately 417 μmol/L (equal to 7 mg/dl). Among the study population, 2,945 individuals with normal serum urate (<417 μmol/L) were treated as healthy controls and 1,387 individuals with high serum urate measured on physical examination were considered hyperuricemia patients. In addition, the 4,332 individuals were also divided into subgroups according to body mass index (BMI) values following the categories of the World Health Organization. Individuals were divided into three BMI subgroups (underweight: BMI < 18.5; normal weight: 18.50 ≤ BMI < 25; overweight: BMI ≥ 25) in this study. Overweight individuals contained those subjects considered obese with BMI greater than 30. The characteristics of the individuals in this study are shown in Table S3 and S4.

RNA Isolation, cDNA Synthesis, and Real-time qPCR. To analyze the cumulative effects of candidate genes in pathway level, we integrated the mRNA expression data of three genes (SLC17A1, SLC17A4 and SLC2A9) that had been measured in our previous study and the transcription levels of the other six candidate genes were analyzed in the present study. All those mRNA expression data expression of candidate genes come from the same samples. Randomly collected RNA from 70 male gout patients, 85 male hyperuricemia patients and 58 male healthy individuals was analyzed. RNA was extracted from blood cells using TRIzol reagent following the manufacturer’s instructions (Invitrogen, Carlsbad, CA, USA). Complementary DNA (cDNA) was synthesized by RNA reverse transcription using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA) following the manufacturer’s protocol. Real-time quantitative polymerase chain reaction (qPCR) was performed using SYBR Premix Ex Taq (TaKaRa Biotech, Tokyo, Japan) with an ABI Prism 7900 Detector System (Applied Biosystems). The data obtained from the assays were analyzed by SDS 2.3 software (Applied Biosystems). The human housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as an internal control.

Target SNPs selection. Based on pathway analysis in GWAS data, potential SNPs for candidate genes in pathway were selected. All potential SNPs were annotated by SNPnexus (http://www.snp-nexus.org/). And then the SNPs filtered by their SNP location and evaluated by the requirements of genotyping technology SNPscan. SNP that located at intergenic region and/or did not satisfy the requirements were replaced by other SNP in the same gene. Finally, after filtration, nine SNPs were chosen from the candidate genes.

Weighted and unweighted genetic urate risk score analysis. To analyze the cumulative effects of the genetic variants in each pathway associated with serum urate, both weighted and unweighted genetic urate risk score (GRS) analysis were applied. In weighted GRS, for each locus, we multiplied the number of effect alleles across the variants in pathways with identical risk. The associations between weight/unweighted GRS and the serum urate concentration in males. The unweighted GRS was calculated by summing the number of effect alleles across the variants in pathways with identical risk. The associations between weight/unweighted GRS and hyperuricemia/gout in males were tested by logistic regression with adjustment for age. A linear regression was used to analyze the relationship between weight/unweighted GRS and the serum urate concentration in males.

Immunoblotting analysis. In the study, five individuals in each of three groups (gout patient, hyperuricemia patient and healthy control) were enrolled from Changhai Hospital in Shanghai and their blood samples were collected to perform Western blotting analysis. All participants signed the written informed consent form. PBMCs were extracted by density-gradient centrifugation (Ficoll-PaqueTM PLUS; GE Healthcare Bio, Chicago, IL, USA). And then the cell lysates from the PBMCs were applied for Western blotting analysis in this study. Equal amounts of protein from samples were subjected to 12% SDS PAGE after which protein was transferred to PVDF membranes (Millipore, Billerica, MA, USA). Membranes were blocked in 5% milk in TBST at room temperature for 2h after which they were incubated with one of the following primary antibodies: rabbit polyclonal to β-actin (1:1000) (Abcam, Hong Kong, Ltd.), and internal control GAPDH (1:3000–5000) (Vazyme, China) at 4°C overnight. Membranes were washed three times with TBST for a total of 30 minutes and then incubated with the horseradish peroxidase-conjugated secondary antibody of goat anti-rabbit, or anti-mouse IgG for 2h at room temperature. The protein bands were visualized with ECL solution.

Statistical analyses. The genotypes at each locus were evaluated for deviation from Hardy-Weinberg equilibrium. Associations with serum urate were tested by linear regression and deviance analysis for linear regression, with adjustments for gender and age. Fisher's exact test in the additive model and logistic regression adjusted
for gender were used to test the association between each locus and gout. Gender subgroups were used in the serum urate and gout analyses. Furthermore, subgroups of body mass index (BMI) was also used in the serum urate analysis. A nominal \( P \) value was calculated and was corrected for multiple testing using the FDR method (\( P_{\text{FDR}} \)). \( P_{\text{FDR}} \) values less than 0.05 were considered statistically significant.

The associations between the pathways and the two phenotypes (serum urate and gout) were tested using the Simes test. The Simes test improves upon the Bonferroni method by rejecting \( H_0 \) for any \( m = 1, 2, \ldots, N \), \( P(\hat{m}) \leq \alpha / N \), where \( P(\hat{m}) \leq \alpha / N \), showing the ordered \( P \) values, with \( P(\hat{m}) \) being the \( m \)th smallest \( P \) value. In our study, the \( P \) values were replaced with \( P_{\text{FDR}} \) values of the loci to control the false positive rate. In addition, the \( P \) value of the pathway was also multiple corrected using the FDR method (\( P_{\text{FDR}} \)).

The mRNA expression data were visualized using box plots, and all outliers were removed. The differences between the different groups in the transcription levels of the candidate genes in the identified pathways were analyzed using the Wilcoxon rank sum test. The differences in gene expression at the pathway level were tested using the Simes test. A \( P \) value less than 0.05 was considered statistically significant.

All statistical analyses were performed using R (Version 3.0.2: www.r-project.org/).

References

1. Marchetti, M. et al. Catalysis and Structure of Zebrafish Urate Oxidase Provide Insights into the Origin of Hyperuricemia in Hominoids. Sci. Rep. 6, 38302 (2016).
2. Kratzer, J. T. et al. Evolutionary history and metabolic insights of ancient mammalian uricases. Proc. Natl. Acad. Sci. USA 111, 3763–8 (2014).
3. Mandal, A. K. & Mount, D. B. The Molecular Physiology of Uric Acid Homeostasis. J. Biol. Chem. 285, 35123–35132 (2010).
4. Yu, F. et al. The Glutathione S-Transferase 1 and Uricase Gene Are Associated with Serum Uric Acid Levels in Members of Sundanese Families. Am. J. Hum. Genet. 83, 673–681 (2008).
5. Uchida, Y. A. et al. Wide expression of type I Na\(^+\)-phosphate cotransporter 3 (NPT3/SLC17A6) in uptake of biotin and organic anion exporter. AJP Cell Physiol. 285, 26107–26113 (2010).
6. Chamley, Y. et al. Type 1 sodium-dependent phosphate transporter (SLC17A1 protein) is a Cl−-dependent urate exporter. J. Neurochem. 67, 281–287 (1996).
7. Saito, M. et al. Type 1 sodium-dependent phosphate transporter (SLC17A1 protein) is a Cl−-dependent urate exporter. J. Biol. Chem. 285, 26107–26113 (2010).
8. Juhatma, P. et al. Human sodium phosphate transporter 4 (hNPT4/SLC17A3) is a common renal secretory pathway for drugs and urate. J. Biol. Chem. 285, 35123–35132 (2010).
9. Tan, A. et al. Genome-wide association study for serum urate concentrations and gout among African Americans identifies genomic risk loci and a novel URA1 loss-of-function allele. Hum. Mol. Genet. 20, 4056–4066 (2011).
36. Lee, P. H. & Shatkay, H. F-SNP: Computationally predicted functional SNPs for disease association studies. *Nucleic Acids Res.* **36** (2008).
37. Zhang, K. *et al.* ICSPNPathway: Identify candidate causal SNPs and pathways from genome-wide association study by one analytical framework. *Nucleic Acids Res.* **39** (2011).
38. Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M. & Hirakawa, M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* **38**, D355–D360 (2010).
39. Ashburner, M. *et al.* Gene Ontology: Tool for The Unification of Biology. *Nat. Genet.* **25**, 25–29 (2000).
40. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–50 (2005).
41. Wallace, S. L. *et al.* Preliminary criteria for the classification of the acute arthritis of primary gout. *Arthritis Rheum.* **20**, 895–900 (1977).
42. Wang, X. *et al.* Rationales, design and recruitment of the Taizhou Longitudinal Study. *BMC Public Health* **9**, 223 (2009).
43. Khanna, D. *et al.* 2012 American college of rheumatology guidelines for management of gout. part 1: Systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. *Arthritis Care Res.* **64**, 1431–1446 (2012).
44. Yang, L. *et al.* Common UCP2 variants contribute to serum urate concentrations and the risk of hyperuricemia. *Sci. Rep.* **6**, 27279 (2016).
45. Dayem Ullah, A. Z., Lemoine, N. R. & Chelala, C. A practical guide for the functional annotation of genetic variations using SNPnexus. *Br. Bioinform* **14**, 437–447 (2013).
46. Simes, R. J. An improved bonferroni procedure for multiple tests of significance. *Biometrika* **73**, 751–754 (1986).

Acknowledgements
The authors thank Fudan University High-End Computing Center for support. This work was supported by grants from National Natural Science Foundation of China (31521003), Ministry of Science and Technology [2012CB944600 to L.J. and F.Z., 2011BAI09B00 to L.J.], the Science and Technology Committee of Shanghai Municipality [11D1400100 to J.W. and H.Z.], International S&T Cooperation Program of China [2013DFA30870 to J.W.], and Program for 2012 Outstanding Medical Academic Leader for Hejian Zou.

Author Contributions
Conceived and designed the experiments: J.W. and Z.D. Performed the experiments: Z.D. Analyzed the data: Z.D. Contributed reagents/materials/analysis tools: J.Z., X.X., S.J., Y.L., D.Z., C.Y., Y.M., Y.W., H.H., H.J., J.Z., Z.Y., Y.Y., X.W. and Y.P. Wrote the paper: J.W., Z.D., H.Z. and L.J.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-21858-0.

Competing Interests: The authors declare 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) 2018