Heritability and Genome-Wide Association Analyses of Serum Uric Acid in Middle and Old-Aged Chinese Twins

Weijing Wang1,2, Dongfeng Zhang1*, Chunsheng Xu1,3,4, Yili Wu1, Haiping Duan1,3, Shuxia Li5 and Qihua Tan2,5
1 Department of Epidemiology and Health Statistics, Public Health College, Qingdao University, Qingdao, China,
2 Epidemiology and Biostatistics, Department of Public Health, University of Southern Denmark, Odense, Denmark, 3 Qingdao Municipal Center for Disease Control and Prevention, Qingdao, China, 4 Qingdao Institute of Preventive Medicine, Qingdao, China, 5 Unit of Human Genetics, Department of Clinical Research, University of Southern Denmark, Odense, Denmark

Serum uric acid (SUA), as the end product of purine metabolism, has proven emerging roles in human disorders. Here based on a sample of 379 middle and old-aged Chinese twin pairs, we aimed to explore the magnitude of genetic impact on SUA variation by performing sex-limitation twin modeling analyses and further detect specific genetic variants related to SUA by conducting a genome-wide association study. Monozygotic (MZ) twin correlation for SUA level \((r_{MZ} = 0.56)\) was larger than for dizygotic (DZ) twin correlation \((r_{DZ} = 0.39)\). The common effects sex-limitation model provided the best fit with additive genetic parameter \((A)\) accounting for 46.3%, common or shared environmental parameter \((C)\) accounting for 26.3% and unique/nonshared environmental parameter \((E)\) accounting for 27.5% for females and 29.9, 33.1, and 37.0% for males, respectively. Although no SUA-related genetic variants reached genome-wide significance level, 25 SNPs were suggestive of association \((P < 1 \times 10^{-5})\). Most of the SNPs were located in an intrinsic region and detected to have regulatory effects on gene transcription. The cell-type specific enhancer of skeletal muscle was detected which has been reported to implicate SUA. Two promising genetic regions on chromosome 17 around rs2253277 and chromosome 14 around rs11621523 were found. Gene-based analysis found 167 genes nominally associated with SUA level \((P < 0.05)\), including \(PTGR2\), \(ENTPD5\), well-known \(SLC2A9\), etc. Enrichment analysis identified one pathway of transmembrane transport of small molecules and 20 GO gene sets involving in ion transport, transmembrane transporter activity, hydrolase activity acting on acid anhydrides, etc. In conclusion, SUA shows moderate heritability in women and low heritability in men in the Chinese population and genetic variations are significantly involved in functional genes and regulatory domains that mediate SUA level. Our findings provide clues to further elucidate molecular physiology of SUA homeostasis and identify new diagnostic biomarkers and therapeutic targets for hyperuricemia and gout.

Keywords: Chinese twins, gene-based test, genome-wide association study, heritability, serum uric acid

Abbreviations: BMI, body mass index; DZ, dizygotic; GSEA, gene set enrichment analysis; GWAS, genome-wide association study; MZ, monozygotic; SUA, serum uric acid; VEGAS2, Versatile Gene-based Association Study-2.
INTRODUCTION

Serum uric acid (SUA), as the end product of purine metabolism, has proven emerging roles in human disorders, such as kidney disease (1, 2), diabetic nephropathy (3), metabolic diseases (4, 5), preeclampsia (6), cardiovascular disease (2, 7), diabetes (8), etc. A systematic review and meta-analysis by Liu et al. concluded that the prevalence of hyperuricemia (13.3%) and gout (1.1%) was high in mainland China (9). Hence, it is necessary to explore factors affecting SUA homeostasis and elucidate underlying pathogenesis of increased SUA level.

The SUA level is mediated by the interplay between genetic and environmental factors. So far, the magnitude of genetic sources of variance in SUA level has been previously explored in several population studies (10–16). A strong genetic component was indicated with heritability estimates approximately ranging from 35 to 77%. Additionally, the genetic epidemiology has presented an enormous impact on the molecular physiology related to SUA homeostasis by genome-wide association study (GWAS), identifying several genetic loci located in key urate transporters such as SLC22A7, SLC22A9, SLC22A11, SLC22A12, ABCG2, etc., and a number of additional intriguing genetic networks (17–19).

Although intensively deployed, no GWAS has, to our knowledge, yet been performed on a sample of middle and old-aged Chinese twins. Chinese population differs in the genetic constitutions and a multitude of life style like dietary habit, work type, and physical activity from other ethnic populations in the world. Genetically related individuals, such as twin pairs, would highly confer increased power in genetic association analysis and efficiently identify genetic variants underlying human complex diseases (20).

Based on a sample of 379 middle and old-aged Chinese twin pairs, we explore the magnitude of genetic impact on SUA variation by performing twin modeling analyses and replicate previous findings on heritability of SUA level and further conduct a GWAS to detect specific genetic variants associated with SUA.

MATERIALS AND METHODS

Participants

The sample collection was carried out through the Qingdao Twin Registry, and details of study recruitment have been described previously (21, 22). Participants who were with gout, systemic lupus erythematosus, eGFR < 60%, or serum creatinine level >1.4 mg/dL were excluded, and incomplete co-twin pairs were also dropped. The final sample consisted of 379 complete twin pairs with a median age of 50 years (95% range: 41–69 years), including 240 monozygotic (MZ) pairs (114 male and 126 female pairs) and 139 dizygotic (DZ) pairs (41 male, 39 female, and 59 opposite-sex pairs).

All co-twin pairs undertook a health examination after a 10–12 h overnight fast and completed a questionnaire. Serum and plasma were separated from blood cells in the field within 30 min and kept frozen at −80°C. The zygosity was determined by using 16 multiple short tandem sequence repeat DNA markers (23, 24). SUA level was measured on the Semi-automatic Analyzer (Hitachi 7600, Japan) and transformed following Blom’s formula for normality.

This study was approved by the Regional Ethics Committee of the Qingdao CDC Institutional Review Boards. Prior written informed consent was achieved for all participants. The ethical principles of the Helsinki Declaration were followed.

Genotyping and Quality Control

DNA samples of 139 DZ pairs were genotyped on the Illumina’s Infinium Omni2.5Exome-8v1.2 BeadChip platform (Illumina, San Diego, CA, USA). Strong quality control was performed using the genome-wide efficient mixed-model association (GEMMA) (25) by removing the SNPs of call rate (<0.98), Hardy–Weinberg Equilibrium ($P < 1 \times 10^{-6}$), locus missing (>0.05), and minor allele frequency (<0.05). Finally, a total of 1,365,181 SNPs was included for subsequent GWAS analysis.

Statistical Analysis

Heritability

Data preparation and descriptive analyses as well as genetic analyses were performed with SPSS version 22.0 and Mx program,1 respectively. Twin pair phenotypic correlations per zygosity were firstly measured by calculating the Pearson’s product-moment correlation coefficients. The higher correlations of MZ than those of DZ twin pairs indicated the genetic effect on individual differences in SUA level.

Then, standard structural equation modeling methods were used for sex-limitation twin modeling based on the classical twin methods. The variation was decomposed into sources of additive genetic (A), common or shared environmental (C), and unique or nonshared environmental (E) parameters. After the general sex-limitation ACE model was firstly fitted, we then fitted its sub-models: the common effects sex-limitation models by setting the male-specific additive genetic effects ($A_m$) and sex-specific common or shared environmental effects ($C_m$ and $C_f$) to 0 and the scalar sex-limitation models by constraining the variance components for females to be equal to a scalar multiple of the variance components for males, respectively.

In order to choose the best fitting model, the likelihood ratio test was applied to compare the performances between the general sex-limitation model and its sub-models. In the likelihood ratio test, twice the difference in the log likelihoods between models was calculated, and change in chi-square against the change in degrees of freedom were tested. The Akaike’s information criterion (AIC) was calculated and a lower AIC indicated a better fit when no statistical difference was observed between two models (26). The covariates of age and body mass index (BMI) were adjusted for in the analysis.

Genome-Wide Association Study

SNPs-Based Genome-Wide Association Study

The association of SUA level with SNP genotypes was tested using the GEMMA (25). The covariates of age, sex, BMI, and the first five principle components were adjusted for in the model fitting.

1http://www.vcu.edu/mx.
The conventional genome-wide significance level of \( P < 5 \times 10^{-8} \) and suggestive evidence level of \( P < 1 \times 10^{-6} \) for this association were adopted (27). We further conducted functional elaboration of GWAS results and predicted putative causal variants in haplotype blocks, likely cell types of action and candidate target genes of noncoding genome by using online HaploReg v4.1 software\(^1\) (28, 29). A set of 25 query SNPs \( (P < 1 \times 10^{-5}) \) was submitted. The enrichments of cell-type enhancers with uncorrected \( P < 0.05 \) were reported.

**Gene-Based Analysis**

We performed gene-based tests on GWAS summary results by using Versatile Gene-based Association Study-2 (VEGAS2) which uses 1,000 genomes data to model SNP correlations across the autosomes and chromosome X (30, 31). In the test, the evidence for association from all SNPs was aggregated within a per gene while correcting for linkage disequilibrium and gene size, and genes showing more signal or strength of association than expected by chance were identified. The SNPs from “1000G East ASIAN Population” were adopted. The significance of over-represented gene sets was determined by Benjamini and Hochberg method corrected \( P \)-value, i.e., false discovery rate (FDR) <0.05.

**Gene Sets-Based Analysis**

A list of significant genes \( (P < 0.05) \) were included to compute the over-represented gene sets in the gene sets-based analysis using the online version of gene set enrichment analysis (GSEA) program\(^2\) (32, 33). Gene sets of Canonical pathways, BioCarta, KEGG, Reactome, GO biological process, and GO molecular function were selected in MSigDB. The significance of over-represented gene sets was determined by Benjamini and Hochberg method corrected \( P \)-value, i.e., false discovery rate (FDR) <0.05.

### RESULTS

**Heritability**

The final sample contained a total of 379 twin pairs (240 MZ and 139 DZ pairs) with a median age of 50 years (95% range: 41–69 years). The median (95% range) of SUA level for all participating individuals was 256 µmol/L (143–468 µmol/L), with males having higher SUA level than females [298 (179–509) vs. 226 (130–374), \( P < 0.001 \)] (Table S1 in Supplementary Material).

After adjusting for the effects of age, sex, and BMI, MZ twin correlation for SUA level (\( r_{MZ} = 0.56, 95\% \) CI: 0.47–0.64) was larger than for DZ twin (\( r_{DZ} = 0.39, 95\% \) CI: 0.25–0.50), indicating the presence of genetic influence (Table S2 in Supplementary Material). As described, we firstly fitted the general sex-limitation (Model II) provided the best fit (AIC = 443.56, \( P > 0.05 \)) with A parameter accounting for 46.3% (95% CI: 15.0–73.4), C parameter accounting for 26.3% (95% CI: 0–55.8), and E parameter accounting for 27.5% (95% CI: 19.8–37.9) for

\[
\begin{align*}
\text{Model I: general sex-limitation model.} \\
\text{Model IIa: sex-limitation model, in which the C \( f \) was set to 0 based on Model II.} \\
\text{Model III: common effects sex-limitation model, in which the C \( f \) was set to 0 based on Model II.} \\
\text{Model IV: common effects sex-limitation model, in which females \( C \) was set to 0 based on Model II.} \\
\text{Model V: scalar sex-limitation model, in which females \( C \) was set to 0 based on Model II.} \\
\text{Model VI: general sex-limitation model, in which the A \( f \) was set to 0 based on Model VI.} \\
\text{Model VII: common effects sex-limitation model, in which females \( A \) was set to 0 based on Model VII.} \\
\text{Model VIII: scalar sex-limitation model, in which females \( A \) was set to 0 based on Model VII.} \\
\end{align*}
\]

\( ^1 \)http://archive.broadinstitute.org/mammals/haploreg/haploreg.php.  
\( ^2 \)http://software.broadinstitute.org/gsea/msigdb/annotate.jsp.
females and 29.9% (95% CI: 0–60.0), 33.1% (95% CI: 5.4–63.4), and 37.0% (95% CI: 26.9–50.3) for males, respectively (Table 1).

**Genome-Wide Association Study**

A total of 1,365,181 SNPs genotyped from a sample of 139 DZ twin pairs were included for the GWAS of SUA level. The relationship between the observed and expected GWAS P-values was illustrated in the Q–Q plot (Figure 1). No evidence of genomic inflation of the test statistics or the bias from the possible population stratification was indicated (λ-statistic = 1). And the slight deviation in the upper right tail from the null distribution suggested evidence for weak association. None of the SNPs reached the genome-wide significance level as illustrated in Figure 2; however, 25 SNPs were suggestive of association ($P < 1 \times 10^{-5}$)

![Figure 1](Quantile–quantile plot for quality control check and visualizing crude association for genome-wide association study of serum uric acid (SUA) level. The x-axis shows the $-\log_{10}$ of expected $P$-values of association from chi-square distribution and the y-axis shows the $-\log_{10}$ of $P$-values from the observed chi-square distribution. The black dots represent the observed data with top hit SNP being colored, and the red line is the expectation under the null hypothesis of no association. Gene at the best SNP is indicated.)

![Figure 2](Manhattan plot for genome-wide association study of serum uric acid (SUA) level. The x-axis shows the numbers of autosomes and the X chromosome, and the y-axis shows the $-\log_{10}$ of $P$-values for statistical significance. The dots represent the SNPs. None of the SNPs reached the genome-wide significance level ($P < 5 \times 10^{-8}$); however, 25 SNPs were suggestive of association ($P < 1 \times 10^{-5}$).)
(Table 2). The strongest association was detected with rs346750 ($P = 2.50 \times 10^{-5}$) in an intronic region of EXOC3L2 on chromosome 19q13.32.

Among these top signals, two chromosomal loci (17q25.3 and 14q24.3) showed nominal association with SUA level as the locus zoom plots illustrated (Figures 3 and 4). On chromosome 17q25.3, seven SNPs ($P = 3.25 \times 10^{-5}$–7.29 $\times 10^{-6}$) and two SNPs ($P = 1.30 \times 10^{-6}$–8.93 $\times 10^{-6}$) were located at or near TNRC6C and TMC6/TNRC6C-AS1 genes, respectively. At chromosome 14q24.3, four SNPs rs2270073, rs2270074, rs16121523, and rs2159179 ($P = 2.77 \times 10^{-5}$–7.59 $\times 10^{-6}$) were positioned within or closest to PTGR2 gene that was involved in terminal inactivation of prostaglandins. The rs2336742 ($P = 7.02 \times 10^{-6}$) was located at the intronic region of ENTPD5 gene, which was involved in the pathway of purine metabolism. The number of SNPs mapping to ZNF410 and FAM161B was five ($P = 3.39 \times 10^{-6}$–7.64 $\times 10^{-6}$) and one ($P = 7.64 \times 10^{-6}$), respectively. All the abovementioned genes showed nominal association with SUA level ($P < 0.05$) from the following VEGAS2 analysis.

As predicted by HaploReg v4.1, two cell-type specific enhancers (uncorrected $P < 0.05$) of brain angular gyrus ($P = 0.003$) and skeletal muscle (female) ($P = 0.008$) were identified for the set of 25 query SNPs (Table S3 in Supplementary Material). Most of the SNPs were located in intronic regions. Several SNPs were detected within regions with promoter histone marks or enhancer histone marks and could change DNA motifs for DNA-binding proteins, and thus would have regulatory effects on gene transcription.

As predicted by HaploReg v4.1, two cell-type specific enhancers (uncorrected $P < 0.05$) of brain angular gyrus ($P = 0.003$) and skeletal muscle (female) ($P = 0.008$) were identified for the set of 25 query SNPs (Table S3 in Supplementary Material). Most of the SNPs were located in intronic regions. Several SNPs were detected within regions with promoter histone marks or enhancer histone marks and could change DNA motifs for DNA-binding proteins, and thus would have regulatory effects on gene transcription.
**DISCUSSION**

In this investigation based on 379 twin pairs, we explored the proportion of genetic sources in SUA variation, and confirmed the genetic variants underlying this trait by GWAS. The MZ twin correlation for SUA level was larger than for DZ twin, indicating the presence of genetic influence (Table S2 in Supplementary Material). In fitting the sex-limitation ACE models, the common effects sex-limitation model of dropping $A_m$ effects (Model II) was favored over the general model (Model I) (AIC = 443.56, $P > 0.05$), indicating that there was no evidence for sex-specific additive genetic effects. We then considered whether the common or shared environmental effects for males ($C_m$) (Model III) and further for females ($C_f$) (Model IV) could also be fixed to 0. However, the goodness-of-fit statistics indicated that these two models provided a significantly worse fit ($P < 0.05$). Thus, these two models were rejected and model II remained the favored one. Finally, we considered the scalar sex-limitation models (Model V and Model VI) by constraining the variance components of females to be equal to scalar multiples of the variance components of males. Both of the models provided a significantly worse fit than Model II ($P < 0.05$). Hence, we concluded that the Model II was the best fitting model, in which additive genetic parameter ($A$) explained larger proportion of SUA variation for females (46.3%) than for males (29.9%), whereas the environmental parameters together ($C$ and $E$) explained smaller proportion (53.7 vs. 70.1%) (Table 1).

The sex-difference in the genetic and environmental effects on SUA variation obtained here was in line with the previous Boyle et al’s genetic study. Based on a sample of 112 twin pairs, they also found a more significant genetic component in control of SUA variation in females than males, whereas a stronger role of environmental component in males (34). We speculated that the genetic architecture may indeed differ across sexes because of the sex differences in selective pressures during human evolution. Additional data from adopted twins and siblings reared together may be used to explore this hypothesis further.

Although no genome-wide significant SNPs were identified in GWAS, we found two promising genetic regions on chromosome 17 around rs2253277 (Figure 3) and chromosome 14 around rs11621523 (Figure 4). The $PTGR2$ and $ENTPD5$ genes around the rs11621523 have been emphasized for their roles in selective pressures during human evolution. Prostaglandin reductase 2 ($PTGR2$) is the enzyme involved in terminal inactivation of prostaglandins (35) which...
FIGURE 4 | Regional association plot showing signal around chromosomal loci of 14q24.3 for genome-wide association study of serum uric acid (SUA) level. The strongest association was detected with rs11621523 in PTGR2 gene.

TABLE 3 | The top 20 genes from Versatile Gene-based Association Study-2 gene-based analysis showing the strongest association with serum uric acid level.

<table>
<thead>
<tr>
<th>Chr</th>
<th>Gene</th>
<th>Number of SNPs</th>
<th>Start position</th>
<th>Stop position</th>
<th>Gene-based test statistic</th>
<th>P-value</th>
<th>Top-SNP</th>
<th>Top-SNP P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>GPR151</td>
<td>3</td>
<td>145,894,416</td>
<td>145,895,676</td>
<td>32.41</td>
<td>5.00E-06</td>
<td>rs11621523</td>
<td>5.30E-05</td>
</tr>
<tr>
<td>17</td>
<td>TNRC6C-AS1*</td>
<td>2</td>
<td>76,103,482</td>
<td>76,107,880</td>
<td>38.89</td>
<td>1.30E-05</td>
<td>rs2748431</td>
<td>8.90E-06</td>
</tr>
<tr>
<td>14</td>
<td>ZNF410*</td>
<td>9</td>
<td>74,353,317</td>
<td>74,398,991</td>
<td>95.92</td>
<td>1.40E-05</td>
<td>rs6574154</td>
<td>5.70E-06</td>
</tr>
<tr>
<td>17</td>
<td>TNRC6C*</td>
<td>30</td>
<td>76,000,317</td>
<td>76,104,916</td>
<td>253.58</td>
<td>2.30E-05</td>
<td>rs1079120</td>
<td>3.00E-06</td>
</tr>
<tr>
<td>5</td>
<td>JMY</td>
<td>14</td>
<td>78,531,924</td>
<td>78,623,038</td>
<td>138.00</td>
<td>4.80E-05</td>
<td>rs2913837</td>
<td>3.00E-05</td>
</tr>
<tr>
<td>14</td>
<td>FAM161B*</td>
<td>11</td>
<td>74,990,694</td>
<td>74,417,117</td>
<td>92.90</td>
<td>4.80E-05</td>
<td>rs1005664</td>
<td>7.60E-06</td>
</tr>
<tr>
<td>17</td>
<td>TNRC6C</td>
<td>30</td>
<td>76,000,317</td>
<td>76,104,916</td>
<td>253.58</td>
<td>2.30E-05</td>
<td>rs1079120</td>
<td>3.00E-06</td>
</tr>
<tr>
<td>5</td>
<td>HOMER1</td>
<td>50</td>
<td>78,669,646</td>
<td>78,809,659</td>
<td>277.24</td>
<td>1.00E-04</td>
<td>rs6794113</td>
<td>3.00E-05</td>
</tr>
<tr>
<td>12</td>
<td>SLC10A1*</td>
<td>98</td>
<td>20,963,637</td>
<td>21,069,843</td>
<td>712.77</td>
<td>1.20E-04</td>
<td>rs1304539</td>
<td>4.40E-05</td>
</tr>
<tr>
<td>14</td>
<td>PTGR2*</td>
<td>11</td>
<td>74,315,533</td>
<td>74,352,168</td>
<td>59.66</td>
<td>2.20E-04</td>
<td>rs2270073</td>
<td>7.40E-06</td>
</tr>
<tr>
<td>10</td>
<td>C10orf32-ASMT</td>
<td>18</td>
<td>104,613,966</td>
<td>104,661,655</td>
<td>109.59</td>
<td>2.60E-04</td>
<td>rs10786719</td>
<td>1.70E-04</td>
</tr>
<tr>
<td>9</td>
<td>LAMC3</td>
<td>88</td>
<td>133,884,503</td>
<td>133,968,446</td>
<td>311.55</td>
<td>3.00E-04</td>
<td>rs10901336</td>
<td>5.40E-04</td>
</tr>
<tr>
<td>14</td>
<td>COX6</td>
<td>11</td>
<td>74,416,636</td>
<td>74,429,813</td>
<td>54.92</td>
<td>3.00E-04</td>
<td>rs4903159</td>
<td>4.00E-05</td>
</tr>
<tr>
<td>10</td>
<td>ASMT</td>
<td>12</td>
<td>104,629,209</td>
<td>104,661,655</td>
<td>72.58</td>
<td>3.20E-04</td>
<td>rs10786719</td>
<td>1.70E-04</td>
</tr>
<tr>
<td>17</td>
<td>CDRT4</td>
<td>18</td>
<td>15,339,331</td>
<td>15,370,925</td>
<td>150.75</td>
<td>3.20E-04</td>
<td>rs76787574</td>
<td>2.50E-06</td>
</tr>
<tr>
<td>1</td>
<td>RP65</td>
<td>3</td>
<td>151,313,115</td>
<td>151,319,769</td>
<td>25.75</td>
<td>3.30E-04</td>
<td>rs1752387</td>
<td>6.70E-04</td>
</tr>
<tr>
<td>10</td>
<td>NANOS1</td>
<td>2</td>
<td>120,789,227</td>
<td>120,793,854</td>
<td>21.94</td>
<td>3.30E-04</td>
<td>rs79664216</td>
<td>8.50E-04</td>
</tr>
<tr>
<td>17</td>
<td>EFNB3</td>
<td>3</td>
<td>7,698,519</td>
<td>7,614,893</td>
<td>33.50</td>
<td>3.40E-04</td>
<td>rs7141</td>
<td>1.40E-04</td>
</tr>
<tr>
<td>10</td>
<td>MYOZ1</td>
<td>4</td>
<td>75,391,369</td>
<td>75,401,515</td>
<td>32.80</td>
<td>3.40E-04</td>
<td>rs1100726</td>
<td>1.10E-04</td>
</tr>
<tr>
<td>10</td>
<td>EIF3A</td>
<td>15</td>
<td>120,794,540</td>
<td>120,840,334</td>
<td>98.21</td>
<td>3.40E-04</td>
<td>rs10787901</td>
<td>8.20E-04</td>
</tr>
<tr>
<td>19</td>
<td>EML2</td>
<td>29</td>
<td>46,112,657</td>
<td>46,148,775</td>
<td>117.77</td>
<td>3.60E-04</td>
<td>rs6509226</td>
<td>1.40E-04</td>
</tr>
</tbody>
</table>

*Represented the genes had already been indicated in the SNPs-based analysis.
### Table 4: The gene sets results—one REACTOME gene set and top 20 GO gene sets (FDR q-value < 0.01) using gene set enrichment analysis (GSEA) program.

<table>
<thead>
<tr>
<th>Gene set name</th>
<th>Genes in gene set (K)</th>
<th>Description</th>
<th>Genes in overlap (k)</th>
<th>k/K</th>
<th>P-value</th>
<th>FDR q-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>REACTOME_TRANSMEMBRANE_TRANSPORT_OF_SMALL_MOLECULES</td>
<td>413</td>
<td>Genes involved in transmembrane transport of small molecules</td>
<td>9</td>
<td>0.0218</td>
<td>1.52E-05</td>
<td>2.02E-02</td>
</tr>
<tr>
<td>GO_MOVEMENT_OF_CELL_OR_SUBCELLULAR_COMPONENT</td>
<td>1,275</td>
<td>The directed, self-propelled movement of a cell or subcellular component without the involvement of an external agent such as a transporter or a pore</td>
<td>18</td>
<td>0.0141</td>
<td>5.12E-07</td>
<td>1.81E-03</td>
</tr>
<tr>
<td>GO PROTEIN.Localization</td>
<td>1,805</td>
<td>Any process in which a protein is transported to, or maintained in, a specific location</td>
<td>21</td>
<td>0.0116</td>
<td>1.25E-06</td>
<td>1.81E-03</td>
</tr>
<tr>
<td>GO_REGULATION_OF_INTRACELLULAR_SIGNAL_TRANSDUCTION</td>
<td>1,656</td>
<td>Any process that modulates the frequency, rate or extent of intracellular signal transduction</td>
<td>20</td>
<td>0.0121</td>
<td>1.30E-06</td>
<td>1.81E-03</td>
</tr>
<tr>
<td>GO_LOCOMOTION</td>
<td>1,114</td>
<td>Self-propelled movement of a cell or organism from one location to another</td>
<td>16</td>
<td>0.0144</td>
<td>1.84E-06</td>
<td>1.81E-03</td>
</tr>
<tr>
<td>GO_REGULATION_OF_MRNA_CATABOLIC_PROCESS</td>
<td>26</td>
<td>Any process that modulates the rate, frequency, or extent of a mRNA catabolic process, the chemical reactions and pathways resulting in the breakdown of RNA, ribonucleic acid, one of the two main type of nucleic acid, consisting of a long, unbranched macromolecule formed from ribonucleotides joined in 3',5'-phosphodiester linkage</td>
<td>4</td>
<td>0.1538</td>
<td>1.94E-06</td>
<td>1.81E-03</td>
</tr>
<tr>
<td>GO_ION_TRANSPORT</td>
<td>1,262</td>
<td>The directed movement of charged atoms or small charged molecules into, out of or within a cell, or between cells, by means of some agent such as a transporter or pore</td>
<td>17</td>
<td>0.0135</td>
<td>2.04E-06</td>
<td>1.81E-03</td>
</tr>
<tr>
<td>GO_BIOLOGICAL_ADHESION</td>
<td>1,032</td>
<td>The attachment of a cell or organism to a substrate, another cell, or other organism. Biological adhesion includes intracellular attachment between membrane regions</td>
<td>15</td>
<td>0.0145</td>
<td>3.39E-06</td>
<td>2.59E-03</td>
</tr>
<tr>
<td>GO_CATION_TRANSPORT</td>
<td>796</td>
<td>The directed movement of cations, atoms, or small molecules with a net positive charge, into, out of or within a cell, or between cells, by means of some agent such as a transporter or pore</td>
<td>13</td>
<td>0.0163</td>
<td>4.55E-06</td>
<td>3.04E-03</td>
</tr>
<tr>
<td>GO_MOTOR_ACTIVITY</td>
<td>131</td>
<td>Catalysis of the generation of force resulting either in movement along a microfilament or microtubule or in torque resulting in membrane scission, coupled to the hydrolysis of a nucleoside triphosphate</td>
<td>6</td>
<td>0.0458</td>
<td>6.83E-06</td>
<td>4.05E-03</td>
</tr>
<tr>
<td>GO_NEUROGENESIS</td>
<td>1,402</td>
<td>Generation of cells within the nervous system</td>
<td>17</td>
<td>0.0121</td>
<td>8.14E-06</td>
<td>4.35E-03</td>
</tr>
<tr>
<td>GO_POSITIVE_REGULATION_OF_CATABOLIC_PROCESS</td>
<td>395</td>
<td>Any process that activates or increases the frequency, rate, or extent of the chemical reactions and pathways resulting in the breakdown of substances</td>
<td>9</td>
<td>0.0228</td>
<td>1.07E-05</td>
<td>5.19E-03</td>
</tr>
<tr>
<td>GO_SMALL_MOLECULE_METABOLIC_PROCESS</td>
<td>1,767</td>
<td>The chemical reactions and pathways involving small molecules, any low molecular weight, monomeric, non-encoded molecule</td>
<td>19</td>
<td>0.0108</td>
<td>1.27E-05</td>
<td>5.33E-03</td>
</tr>
<tr>
<td>GO_REGULATION_OF_CELL_ADHESION</td>
<td>629</td>
<td>Any process that modulates the frequency, rate or extent of attachment of a cell to another cell or to the extracellular matrix.</td>
<td>11</td>
<td>0.0175</td>
<td>1.35E-05</td>
<td>5.33E-03</td>
</tr>
<tr>
<td>GO_REGULATION_OF_OF_NUCLEAR_TRANSCRIBED_MRNA_CATABOLIC_PROCESS_DEADENYLATION_DEPENDENT_DECAY</td>
<td>15</td>
<td>Any process that modulates the frequency, rate or extent of nuclear-transcribed mRNA catabolic process, deadenylation-dependent decay</td>
<td>3</td>
<td>0.2</td>
<td>1.79E-05</td>
<td>6.14E-03</td>
</tr>
<tr>
<td>GO_SECONDARY_ACTIVE_TRANSMEMBRANE_TRANSPORTER_ACTIVITY</td>
<td>233</td>
<td>Catalysis of the transfer of a solute from one side of a membrane to the other, up its concentration gradient. The transporter binds the solute and undergoes a series of conformational changes. Transport works equally well in either direction and is driven by a chemiosmotic source of energy, not direct ATP coupling. Chemiosmotic sources of energy include uniport, symport, or antiport</td>
<td>7</td>
<td>0.03</td>
<td>1.83E-05</td>
<td>6.14E-03</td>
</tr>
<tr>
<td>GO_POSITIVE_REGULATION_OF_MRNA_METABOLIC_PROCESS</td>
<td>45</td>
<td>Any process that activates or increases the frequency, rate, or extent of mRNA metabolic process</td>
<td>4</td>
<td>0.0889</td>
<td>1.84E-05</td>
<td>6.14E-03</td>
</tr>
</tbody>
</table>
may contribute to renal uric acid metabolism (36, 37). Four highly correlated SNPs showing suggestive evidence of association with SUA level were detected within or near PTGR2 gene. The rs2270073 and rs2270074 were detected within a region with promoter histone marks in the 5'-UTR of PTGR2 gene and could change DNA motifs for DNA-binding proteins, which provided strong evidence for their regulatory effects on gene transcription (Table S4 in Supplementary Material). Additionally, rs1621523 and rs2159179, which were located in an intergenic region at 14q24.3 and closest to PTGR2 gene, should also be candidates to be further studied. The protein encoding by ENTPD5 gene was involved in the pathway of purine metabolism in which SUA was a by-product of oxidation. As SNP rs2336742 was located at the intronic region of ENTPD5 gene and could change its DNA motifs, it might be associated with purine metabolism as well as SUA level (Table S4 in Supplementary Material). However, the association of novel TNRC6C, TMC6, and TNRC6C-AS1 genes around the other SNP rs225327 with SUA level still needs to be validated. Finally, the enhancer of skeletal muscle was predicted by submitting the set of 25 query SNPs to HaploReg v4.1 (Table S3 in Supplementary Material). And the relationship between SUA level and skeletal muscle strength/volume has been fully researched currently (38–41).

As additional replication, we compared the SUA-associated SNPs reported in a series of studies with ours (18, 42–59) (Table S5 in Supplementary Material). A list of SNPs could be replicated, especially the variants located in the well-known SUA-associated genes SLC2A9 and ABCG2. Notably, two SNPs rs2231142 (ABCG2) and rs10008015 (TET2) were also found by Yang et al. (58) and three SNPs rs11996526 (LOC107986971), rs4848700 (TFCP2L1), and rs179785 (KCNQ1) by Li et al. (49) in Chinese populations.

Table 4: Continued

<table>
<thead>
<tr>
<th>Gene set name</th>
<th>Genes in gene set (K)</th>
<th>Description</th>
<th>Genes in overlap (k)</th>
<th>k/K</th>
<th>P-value</th>
<th>FDR q-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO_Enzyme-linked_recepto...</td>
<td>10</td>
<td>0.0144</td>
<td>3.74E-05</td>
<td>9.01E-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO_Metal_ion_transport</td>
<td>802</td>
<td>Catalysis of the hydrolysis of any acid anhydride</td>
<td>12</td>
<td>0.0146</td>
<td>3.14E-05</td>
<td>8.97E-03</td>
</tr>
<tr>
<td>GO_regulation_of_immune...</td>
<td>1,403</td>
<td>Any process that modulates the frequency, rate, or extent of an immune system process</td>
<td>16</td>
<td>0.0114</td>
<td>3.19E-05</td>
<td>8.97E-03</td>
</tr>
<tr>
<td>GO_Intracellular_signal_transduction</td>
<td>1,572</td>
<td>The process in which a signal is passed on to downstream components within the cell, which become activated themselves to further propagate the signal and finally trigger a change in the function or state of the cell</td>
<td>17</td>
<td>0.0108</td>
<td>3.48E-05</td>
<td>9.01E-03</td>
</tr>
<tr>
<td>GO_Active_transmembrane_transporter_activity</td>
<td>356</td>
<td>Catalysis of the transfer of a specific substance or related group of substances from one side of a membrane to the other, up the solute’s concentration gradient. The transporter binds the solute and undergoes a series of conformational changes. Transport works equally well in either direction</td>
<td>8</td>
<td>0.0225</td>
<td>3.68E-05</td>
<td>9.01E-03</td>
</tr>
<tr>
<td>GO_Cell_motility</td>
<td>835</td>
<td>Any process involved in the controlled self-propelled movement of a cell that results in translocation of the cell from one place to another</td>
<td>12</td>
<td>0.0144</td>
<td>3.74E-05</td>
<td>9.01E-03</td>
</tr>
<tr>
<td>GO_Metal_ion_transport</td>
<td>582</td>
<td>The directed movement of metal ions, any metal ion with an electric charge, into, out of or within a cell, or between cells, by means of some agent such as a transporter or pore</td>
<td>10</td>
<td>0.0172</td>
<td>3.88E-05</td>
<td>9.01E-03</td>
</tr>
</tbody>
</table>

Collections included CP, CP: BIOCARTA, CP: KEGG, and CP: REACTOME of C2: curated gene sets, and BP and MF of C5: GO gene sets in GSEA. Genes in Overlap (k): the number of genes in the intersection of the query set with a set from MSigDB. K: the number of genes in the set from MSigDB. FDR q-value: the false discovery rate analog of hypergeometric P-value after correction for multiple hypothesis testing according to Benjamini and Hochberg.
The association of well-known urate transporter gene SLC2A9 ($P = 0.001$) with SUA level has previously been reported (61, 62). Other genes were of unknown function in terms of SUA level or purine metabolism currently, whereas they may also be interesting potential candidates to be future researched and validated, especially the top 20 genes (Table 3).

Several strengths must be noticed in our study. First, our results based on the twin data of SUA level would be credible. Phenotype variation may be under the effect of subjects’ genetic background, age, gender, and environmental exposures as well as some experimental variables related to sampling, processing, and data analysis. Genetically related individuals, such as twin pairs, would highly confer increased power in genetic association analysis and efficiently identify genetic variants underlying human complex diseases (20). Second, given the various genetic constitutions and multitude of life style among different ethnic populations in the world, this is the first GWAS conducted in the sample of middle and old-aged Chinese twins.

Nevertheless, our study has potential limitations as well. First, our study was with relatively small sample size and limited statistical power resulting from the challenges of identifying and recruiting qualified twin pairs. The results presented here, however, provided a useful reference for hypotheses to be further replicated and validated for exploring increased SUA level. Given the genetic effect on SUA variation is expected to comprise a large number of SNPs possessing very small effect size, a meta-analysis with larger samples will be desirable and ideal. Second, even though we replicated parts of our GWAS results by comparing with results generated from external and independent datasets, most of the SNPs did not reach the genome-wide significance level. In addition, as no other study has explored the differential expressed genes based on SUA-discordant samples, we cannot yet validate our findings further.

In summary, we have confirmed that genetic factors are significant in explaining SUA level variability through twin modeling. Two novel suggestive regions located at chromosomes 17 and 14 were identified. Twenty-five SNPs reached suggestive evidence level of association with SUA and most of them could have regulatory effects on gene transcription, and 167 genes nominally associated with SUA level were involved in significant biological functions related to uric acid generating and metabolism. The potential candidate biomarkers of SUA level reported here should merit further verifications.

REFERENCES


26. Wang et al. Heritability and GWAS of SUA


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Wang, Zhang, Xu, Wu, Duan, Li and Tan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.