Genome-wide association study identifies novel loci associated with serum level of vitamin B12 in Chinese men

Xiaoling Lin1,2,†, Daru Lu3,†, Yong Gao2,4,†, Sha Tao10,†, Xiaobo Yang4,5,†, Junjie Feng11, Aihua Tan4, Haiying Zhang4,5, Yanling Hu4,6, Xue Qin4,7, Seong-Tae Kim11, Tao Peng8,12, Li Li6, Linjian Mo9, Shijun Zhang4, Jeffrey M. Trent10, Zengnan Mo4,9, * S. Lilly Zheng11, ‡ and Jielin Sun11,‡

1Fudan Institute of Urology, Huashan Hospital, 2Fudan-VARI Center for Genetic Epidemiology, School of Life Sciences and 3State Key Laboratory of Genetic Engineering, Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai, China, 4Center for Genomic and Personalized Medicine, 5Department of Occupational Health and Environmental Health, School of Public Health, 6Medical Scientific Research Center, 7Department of Clinical Laboratory, The First Affiliated Hospital, 8Department of Hepatobiliary Surgery, The First Affiliated Hospital, and 9Institute of Urology and Nephrology, The First Affiliated Hospital, Guangxi Medical University, Nanning, Guangxi, China, and 10Center for Genetic Epidemiology, Van Andel Research Institute, Grand Rapids, MI, USA, 11Center for Cancer Genomics, Wake Forest University School of Medicine, Winston-Salem, NC, USA and 12Laboratory of Genomic Diversity, National Cancer Institute, NIH, Frederick, MD, USA

Received September 28, 2011; Revised February 13, 2012; Accepted February 20, 2012

Vitamin B12 (VitB12 or cobalamin) is an essential cofactor in several metabolic pathways. Clinically, VitB12 deficiency is associated with pernicious anemia, neurodegenerative disorder, cardiovascular disease and gastrointestinal disease. Although previous genome-wide association studies (GWAS) identified several genes, including FUT2, CUBN, TCN1 and MUT, that may influence VitB12 levels in European populations, common genetic determinants of VitB12 remain largely unknown, especially in Asian populations. Here we performed a GWAS in 1999 healthy Chinese men and replicated the top findings in an independent Chinese sample with 1496 subjects. We identified four novel genomic loci that were significantly associated with serum level of VitB12 at a genome-wide significance level of $5.00 \times 10^{-8}$. These four loci were MS4A3 (11q12.1; rs2298585; $P = 2.64 \times 10^{-15}$), CLYBL (13q32; rs41281112; $P = 9.23 \times 10^{-10}$), FUT6 (19p13.3; rs3760776; $P = 3.68 \times 10^{-13}$) and 5q32 region (rs10515552; $P = 3.94 \times 10^{-8}$). In addition, we also confirmed the association with the serum level of VitB12 for the previously reported FUT2 gene and identified one novel non-synonymous single-nucleotide polymorphism in FUT2 gene in this Chinese population (19q13.33; rs1047781; $P = 3.62 \times 10^{-36}$). The new loci identified offer new insights into the biochemical pathways involved in determining the serum level of VitB12 and provide opportunities to better delineate the role of VitB12 in health and disease.

INTRODUCTION

Vitamin B12 (VitB12), also called cobalamin, is a water-soluble vitamin. It serves as a cofactor for two enzymes, adenosylcobalamin-dependent methylmalonyl-CoA mutase in mitochondria and methylcobalamin-dependent methionine synthase in the cytoplasm, that play key roles in a range of biological processes including DNA synthesis and regulation.

†The authors wish it to be known that, in their opinion, the first five authors should be regarded as joint First Authors.
‡The authors wish it to be known that, in their opinion, the last three authors should be regarded as joint Senior Authors.

© The Author 2012. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
energy production and the formation of the red blood cells. Deficiency in VitB12 is clinically associated with pernicious anemia, neuropsychiatric symptoms and glossitis (1). VitB12 cannot be synthesized in the human body. Most of the diseases due to VitB12 deficiency were related to poor VitB12 absorption, rather than direct dietary deficiency. VitB12 absorption requires the glycoprotein intrinsic factor from the gastric cell in a functional gastrointestinal (GI) absorption system (2,3). Any factors influencing this system may cause VitB12 malabsorption.

Although the serum VitB12 assay is the best first-line test, the results must be carefully interpreted in clinical context, since a normal value does not preclude the possibility of VitB12 deficiency. Serum level of VitB12 is also influenced by environmental factors, such as age and the intake of folic acid, and at least in part genetically determined. Studying genetic variants which affect VitB12 level is important to understand the interaction among nutrition, genetics and health. Recently, a genome-wide association study (GWAS) in 2717 women of European ancestry identified a strong association between rs492602 in FUT2 and serum VitB12 levels from the Cancer Genetic Markers of Susceptibility projects and the Nurses’ Health Study (NHS-CGEMS) (4), with a P-value of $5.36 \times 10^{-17}$. This single-nucleotide polymorphism (SNP) was in strong linkage disequilibrium (LD) with FUT2 $W143X$ (rs601338) ($r^2 = 0.76$), which determined the FUT2 secretor status. Women who are homozygous for the G-allele of SNP rs492602 had higher VitB12 levels (4). To identify additional genetic factors associated with VitB12 levels, Hazra et al. (5) conducted a meta-analysis of three GWASs based on a total number of 4763 women, including 1658 Caucasian women in NHS-CGEMS, 1647 women in Framingham-SNP-Health Association Resource (SHARE) and 1458 men in SHARE. The meta-analysis identified three additional loci that were associated with VitB12 levels. At 5q32, 9p21.1, 11q12.1, 13q32.3, 19p13.2, 19p13.3 and 19q13.3. We evaluated one SNP from each of these eight genomic regions (Fig. 1). The regions were 1q42.2, 10q24-11.2, 11q12.1, 13q32.3, 19p13.2, 19p13.3 and 19q13.3. We evaluated one SNP from each of these eight regions, which represented the only significant SNP after adjusting for other significant SNPs in each region, in an additional independent set of healthy Chinese men of second stage. The eight SNPs that were evaluated in the second stage were rs583228 at 1q42.2, rs10515552 at 5q32, rs12377462 at 9p21.1, rs2298585 at 11q12.1, rs41281112 at 13q32.3, rs2340550 at 19p13.2, rs3760776 at 19p13.3 and rs1047781 at 19q13.3. In the second stage, four SNPs (rs10515552, rs2298585, rs3760776 and rs1047781) were confirmed to be significantly associated with VitB12 levels at a $P$-value cutoff of 0.006 (Bonferroni correction of eight tests). When the two stages were combined, each of those four SNPs was highly significant, with a $P$-value < $5.00 \times 10^{-8}$, ranging from $3.94 \times 10^{-8}$ to $3.62 \times 10^{-36}$. Although the $P$-value for rs41281112 was slightly higher than the Bonferroni $P$-value cutoff of 0.006 ($P = 0.007$), the combined $P$-value for rs41281112 reached a genome-wide significance level of $5.00 \times 10^{-8}$ ($P$-combined = $9.23 \times 10^{-10}$). The remaining SNPs (rs583228, rs12377462 and rs2340550) were not replicated in the second stage of data (all $P > 0.05$) (Table 2).

### Table 1. General characteristic of the two-stage GWAS study subjects

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>First stage</th>
<th>Second stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1999</td>
<td>1496</td>
</tr>
<tr>
<td>Age (mean ± SD)</td>
<td>37.5 ± 11.1</td>
<td>37.3 ± 10.8</td>
</tr>
<tr>
<td>BMI* (mean ± SD)</td>
<td>23.3 ± 3.4</td>
<td>23.5 ± 3.5</td>
</tr>
<tr>
<td>Folic acid* (mean ± SD)</td>
<td>9.6 ± 2.8</td>
<td>8.3 ± 3.0</td>
</tr>
<tr>
<td>Vitamin B12* (mean ± SD)</td>
<td>697 ± 241</td>
<td>728 ± 297</td>
</tr>
<tr>
<td>Smoking, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1015 (50.8)</td>
<td>771 (51.5)</td>
</tr>
<tr>
<td>No</td>
<td>984 (49.2)</td>
<td>725 (48.5)</td>
</tr>
<tr>
<td>Alcohol drinking, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1704 (85.5%)</td>
<td>1165 (82.6%)</td>
</tr>
<tr>
<td>No</td>
<td>288 (14.5%)</td>
<td>246 (17.4%)</td>
</tr>
</tbody>
</table>

*BMI measured in kg m$^{-2}$; data were log-transformed for GWAS analysis.
*BMI measured in ng/ml; data were log-transformed for GWAS analysis.
*Folic acid measured in pg/ml.

### RESULTS

Demographic information, including age, body mass index (BMI), the distributions of folic acid and VitB12, smoking status and alcohol drinking in the study population are described in Table 1. No significant differences were found between samples of the first and the second stages (Table 1) regarding the mean age, smoking distribution and mean BMI ($P > 0.05$).

In the first stage, no population structure was observed (Supplementary Material, Fig. S1). In addition, we did not observe any evidence for systematic bias of the association for VitB12 phenotype, as indicated by the quantile–quantile (Q–Q) plots (Supplementary Material, Fig. S2). The inflation factor was estimated to be 1.02.

In the first stage of GWASs, we found that multiple SNPs were associated with VitB12 levels at $P < 1.00 \times 10^{-5}$ in eight genomic regions (Fig. 1). The regions were 1q42.2, 5q32, 9p21.1, 11q12.1, 13q32.3, 19p13.2, 19p13.3 and 19q13.3. We evaluated one SNP from each of these eight regions, which represented the only significant SNP after adjusting for other significant SNPs in each region, in an additional independent set of healthy Chinese men of second stage. The eight SNPs that were evaluated in the second stage were rs583228 at 1q42.2, rs10515552 at 5q32, rs12377462 at 9p21.1, rs2298585 at 11q12.1, rs41281112 at 13q32.3, rs2340550 at 19p13.2, rs3760776 at 19p13.3 and rs1047781 at 19q13.3. In the second stage, four SNPs (rs10515552, rs2298585, rs3760776 and rs1047781) were confirmed to be significantly associated with VitB12 levels at a $P$-value cutoff of 0.006 (Bonferroni correction of eight tests). When the two stages were combined, each of those four SNPs was highly significant, with a $P$-value < $5.00 \times 10^{-8}$, ranging from $3.94 \times 10^{-8}$ to $3.62 \times 10^{-36}$. Although the $P$-value for rs41281112 was slightly higher than the Bonferroni $P$-value cutoff of 0.006 ($P = 0.007$), the combined $P$-value for rs41281112 reached a genome-wide significance level of $5.00 \times 10^{-8}$ ($P$-combined = $9.23 \times 10^{-10}$). The remaining SNPs (rs583228, rs12377462 and rs2340550) were not replicated in the second stage of data (all $P > 0.05$) (Table 2).
The first novel VitB12 level associated with locus was located at 11q12.1. Multiple SNPs in an ~170 kb segment (between 59,580 and 59,750 kb) were associated with VitB12 levels with \( P < 1.0 \times 10^{-5} \) in the first stage and included three genes (MS4A3, MS4A2 and MS4A6A) (Supplementary Material, Fig. S3). The SNP rs2298585, in the intron of the MS4A3 gene, provided the strongest statistical evidence \( (P = 1.71 \times 10^{-10}) \) in the first stage and was confirmed in the second stage \( (P = 1.58 \times 10^{-10}) \). The remaining SNPs in the region were no longer significant after adjusting for rs2298585 \( (P > 0.05 \text{ in the first stage}) \). The combined \( P \)-value for rs2298585 was \( 2.64 \times 10^{-15} \). Participants with homozygous rs2298585 ‘T’ alleles had higher VitB12 levels compared with allele ‘C’ carriers \( (P\text{-combined} = 2.64 \times 10^{-15}) \).

The second novel VitB12 level associated with locus was located at 19p13.3. Multiple SNPs in an ~40 kb segment (between 57,700 and 58,100 kb) were associated with VitB12 levels with \( P < 1.0 \times 10^{-5} \) in the first stage, which include two genes (FUT6 and FUT3) (Supplementary Material, Fig. S4). The SNP rs3760776, in the promoter of the FUT6 gene, provided the strongest statistical evidence \( (P = 4.23 \times 10^{-10}) \) in the first stage. The remaining SNPs in the region were no longer significant after adjusting for rs3760776 \( (P > 0.05 \text{ in the first stage}) \). The combined \( P \)-value for rs3760776 was \( 2.64 \times 10^{-15} \). Participants with homozygous rs3760776 ‘T’ alleles had higher VitB12 levels compared with allele ‘C’ carriers \( (P\text{-combined} = 2.64 \times 10^{-15}) \).

The third novel VitB12 level associated with locus is mapped to 5q32 at an intergenetic region, which is located downstream of 3’ of the PRELID2 gene. Multiple SNPs in an ~250 kb segment (between 144,900 and 145,150 kb) were associated with VitB12 levels at \( P < 1.0 \times 10^{-5} \) in the first stage (Supplementary Material, Fig. S5). The SNP rs10515552 provided the strongest statistical evidence \( (P = 8.52 \times 10^{-7}) \) in the first stage. The remaining SNPs in the region were no longer significant after adjusting for rs10515552 \( (P > 0.05 \text{ in the first stage}) \). rs10515552 was tested for replication in the second stage and was confirmed with a \( P \)-value of \( 5.13 \times 10^{-3} \). Men carrying a ‘C’ allele have higher VitB12 levels compared with ‘T’ allele carriers \( (P\text{-combined} = 3.94 \times 10^{-8}) \).

The fourth novel region is located at 19q13.3. Only one SNP (rs41281112) provided strong evidence of association with VitB12 levels in this ~100 kb segment (between 99,256 and 99,356 kb) for VitB12 levels in the first stage (Supplementary Material, Fig. S6). The association of rs41281112 with VitB12 levels was confirmed in the second stage, with a \( P \)-value of \( 9.23 \times 10^{-10} \). Although it did not reach the Bonferroni-corrected \( P \)-value of \( 0.006 \) (assuming eight tests in the second stage), the combined two stage \( P \)-value for rs41281112 reached a genome-wide significant level of \( 5 \times 10^{-8} \) \( (P\text{-combined} = 9.23 \times 10^{-10}) \). The SNP rs41281112 is a non-synonymous SNP found in CLYBL. The substitution of ‘G’ to ‘A’ allele leads to an amino acid change Arg to OPA, a stop codon at amino acid position of 259. Men who carry a homozygous ‘G’ allele for rs41281112 had higher VitB12 levels compared with ‘A’ allele carriers.
Table 2. Results for SNPs associated with VitB12 level from two-stage GWAS study in Chinese population

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr (position)</th>
<th>Genes</th>
<th>Gene regions</th>
<th>Allele (a/A)</th>
<th>First stage (N = 1999)</th>
<th>Second stage (N = 3495)</th>
<th>Combined (N = 3495)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean levels (pg/ml)</td>
<td>Mean levels (pg/ml)</td>
<td>P-value combined</td>
</tr>
<tr>
<td>rsl0515552</td>
<td>5q32</td>
<td>MS4A3</td>
<td>Intergenic</td>
<td>T/C</td>
<td>0.22</td>
<td>679 719</td>
<td>7.68 4.82 ×10^-7</td>
</tr>
<tr>
<td>rsl2298585</td>
<td>13q32.3</td>
<td>MS4A3</td>
<td>Intron</td>
<td>A/G</td>
<td>0.12</td>
<td>681 761</td>
<td>3.34 1.71 ×10^-8</td>
</tr>
<tr>
<td>rs1047781</td>
<td>11q12.1</td>
<td>MS4A3</td>
<td>Intron</td>
<td>A/G</td>
<td>0.12</td>
<td>681 761</td>
<td>3.34 1.71 ×10^-8</td>
</tr>
<tr>
<td>rs3760776</td>
<td>19p13.3</td>
<td>FUT6</td>
<td>Promoter</td>
<td>A/G</td>
<td>0.12</td>
<td>681 761</td>
<td>3.34 1.71 ×10^-8</td>
</tr>
<tr>
<td>rs10515552</td>
<td>5q32</td>
<td>MS4A3</td>
<td>Intergenic</td>
<td>T/C</td>
<td>0.22</td>
<td>679 719</td>
<td>7.68 4.82 ×10^-7</td>
</tr>
<tr>
<td>rs2298585</td>
<td>13q32.3</td>
<td>MS4A3</td>
<td>Intron</td>
<td>A/G</td>
<td>0.12</td>
<td>681 761</td>
<td>3.34 1.71 ×10^-8</td>
</tr>
<tr>
<td>rs1047781</td>
<td>11q12.1</td>
<td>MS4A3</td>
<td>Intron</td>
<td>A/G</td>
<td>0.12</td>
<td>681 761</td>
<td>3.34 1.71 ×10^-8</td>
</tr>
<tr>
<td>rs3760776</td>
<td>19p13.3</td>
<td>FUT6</td>
<td>Promoter</td>
<td>A/G</td>
<td>0.12</td>
<td>681 761</td>
<td>3.34 1.71 ×10^-8</td>
</tr>
</tbody>
</table>

Combined P-values based on multivariate linear regression analysis, adjusted for age, logL-folic acid, BMI and stage information. A indicates serum levels of VitB12 for homozygous carriers of major alleles, Aa indicates heterozygous carriers and aa indicates homozygous carriers of minor alleles.

We also confirmed the association of previously reported FUT2 gene with a serum level of VitB12. The SNP rs1047781, which is a nonsense mutation (Ile140Phe), was the strongest associated SNP with VitB12 levels in an ~100 kb segment (between 53,841 and 53,941 kb) in the first stage (Supplementary Material, Fig. S7). The association of rs1047781 was confirmed in the second stage (P = 6.79 × 10^-22 in the second stage, P-combined = 3.62 × 10^-36).

SNPs rs602662, rs492602 and rs601338 (W143X) found in the FUT2 gene have been reported to be associated with VitB12 levels in two independent GWASs (4.6). However, we did not confirm any of them in our first stage (P > 0.05).

We did find that men who are homozygous for the rs1047781 'T' allele were more likely to have higher VitB12 concentration compared with variant allele 'A' carriers.

We then evaluated the proportion of total variance of VitB12 levels that can be explained by the five SNPs identified in our study: rs10515552 at 5q32, rs2298585 at 11q12.1, rs41281112 at 13q32.3, rs3760776 at 19p13.3 and rs1047781 at 19q13.33, explained 1.1, 1.7, 1.1, 1.4 and 4.0% of the proportion of total variance of VitB12 levels, respectively. In total, these five SNPs account for a higher proportion of variance (9.3%), compared with the variance that can be explained by the other covariates, including age, BMI and folate acid (4.5%). Specifically, only 0.6, 0.3 and 3.6% of total variance can be explained by those three covariates, respectively.

In addition, we examined the association with VitB12 levels for the other three previously reported genes in our stage 1 samples (TCN1, MUT and CUBN) (5). The association results for these SNPs in our first-stage samples are presented in Table 3. Variants in TCN1 (rs526934, P = 0.0007), MUT (rs9473555, P = 0.0004) and CUBN (rs12243895, P = 0.007) were significantly associated with VitB12 levels (Table 3) and were in the same direction of association with previous report (5).

DISCUSSION

To our knowledge, our study represents the first GWAS performed in a Chinese population to search for genetic loci associated with serum levels of VitB12. Using a two-stage design with a total sample size of 3495 healthy men, we successfully identified four novel loci that were significantly associated with VitB12 level, MS4A4A (11q12.1; rs2298585; P = 2.64 × 10^-12), CLYBL (13q32; rs41281112; P = 9.23 × 10^-10), FUT6 (19p13.3; rs3760776; P = 3.68 × 10^-11) and 5q32 region (rs10515552; P = 3.94 × 10^-7). In addition, we also identified one novel non-synonymous SNP in the previously reported FUT2 gene (19q13.33; rs1047781; P = 3.62 × 10^-36).

The most significant novel SNP (rs2298585) identified was located on an ~580 kb region of 11q12.1 of MS4A gene cluster (MS4A3, MS4A2, MS4A6A, MS4A4A and MS4A6E; Supplementary Material, Fig. S3). rs2298585 is located in the intron of MS4A4A gene. MS4A4A encodes a member of the membrane-spanning 4A gene family. The MS4A protein group, which has a characteristic four-transmembrane structure, likely functions as surface signaling molecules and...
intracellular adapter proteins (7–9). On the basis of limited functional studies on MS4A3, it appears that the MS4A3 protein (also called HTm4) may serve as a cell-cycle regulator. In hematopoietic cells, MS4A3 has been shown to inhibit the G1–S cell-cycle transition by suppressing the activation of cyclin-dependent kinase 2 (CDK2) towards cyclin A through its direct stimulation of the CDK2 phosphatase and inhibitor CDKN3/KAP (10,11).

One distinctive feature of VitB12 is that it cannot be synthesized in the human body but instead comes primarily from dietary intake through GI tract absorption, thus the proper maintenance of the GI system is essential for the sufficient absorption of VitB12. Certain pathological conditions, such as malnutrition and certain genetic polymorphisms, may participate in the cell-cycle regulation of GI epithelial cells and that genetic polymorphisms (rs2298585, or SNPs in LD with it) within MS4A3 may interfere with its proper cell-cycle-regulatory role and affect the self-renewal of GI tract which in turn may influence the absorption of VitB12. It should be mentioned, however, that expression of MS4A3 was regarded to be restricted to the hematopoietic and central nervous system in mice (14). Whether or not it is expressed in the GI tract of humans is still not clear and its cell-cycle-regulating role in such a context may also deserve investigation.

In addition, we note that rs2298585 is located ~200–350 kb upstream of transcobalamin I (TCNI) precursor and gastric intrinsic factor (GIF), which encode two of the three major types of VitB12-binding proteins in human (transcobalamin I and III, respectively). The gastric mucosa-secreted GIF promotes absorption of VitB12, and TCNI protein facilitates transport of VitB12 into cells. There are reports that causally correlate TCN1/GIF deficiencies with low serum VitB12 levels and in turn modify the risk of malabsorption of VitB12.

The second novel loci identified was located in the FUT6 gene at chr19q13.33. FUT6 encodes fucosyltransferase, which is a Golgi stack membrane protein involved in the formation of sialyl-Lewis X, an E-selectin ligand. It is noted that absorption of VitB12 is significantly dependent on the GI absorption system. The release of VitB12 requires the secretion of intrinsic factor (IF) from the gastric cell and the formation a VitB12-IF complex. Only the complex could be allowed to enter the gastrointestinal mucosa through receptor-mediated endocytosis. The expression level of Lewis was found to be a determinant for the bacterial density of Helicobacter pylori (19,20). Overgrowth of gastric bacteria, such as H. pylori, can reduce the secretion of IF by reducing the secretion of gastric juice and has a relationship with VitB12 deficiency (21). The SNP rs3760776, which is located in the promoter region of FUT6 may affect the activity of FUT6 enzyme, thus may alter individual’s susceptibility to bacterial infection and in turn modify the risk of malabsorption of VitB12.

The third novel region was mapped to 5q32 at an intergenic gene region, which is 99 kb downstream of PRELID2 gene. PRELID2 belongs to the PREL domain containing family and has been identified as a conserved gene across species. However, the biological relationship between this region and VitB12 levels is unclear. It is also possible that PRELID2 may not be the gene underlying the signal on chromosome 5 and other potential functional regulatory element such as enhancers may contribute to the association on this genomic region.

The fourth novel region was located on the CLYBL gene at chromosome 13q32. CLYBL encodes citrate lyase beta-like protein. The molecular functions of CLYBL include citrate (pro-3s)-lyase activity, carbon–carbon lyase activity and metal ion binding. The substitution of G to A allele of rs41281112 results in a stop codon. Men carrier AA genotypes had significantly lower level of VitB12, compared with men with AG or GG genotypes. On the basis of the above evidence, we hypothesized that the G to A allele substitution of rs41281112 leads to the early termination of translation of CLYBL protein, which may affect normal functioning of CLYBL protein of metal ion binding and may interfere with ion uptake. This may in turn lead to the malabsorption of VitB12.
For the \textit{FUT2} gene, three SNPs, including rs602662, rs492602 and rs601338, were previously reported to be associated with VitB12 levels (4). However, these three SNPs were not confirmed in our first stage. In contrast, another non-synonymous SNP in the \textit{FUT2} gene (rs10447781) was strongly associated with VitB12 levels, with a combined \( P \)-value of \( 1.63 \times 10^{-35} \). The SNP was in weak linkage disequilibrium (\( r^2 = 0.001 \)) with any of the above three reported SNPs in the Chinese population. In addition, rs1047781 was not polymorphic in Europeans. The substitution of A to T allele leads to an amino acid change from isoleucine to phenylalanine in Asian populations. A previous functional study reported that this substitution inactivates fucosyltransferase (22). This evidence indicates that rs1047781 is plausibly the potential functional variant that is associated with VitB12 level in the Asian population.

There were four VitB12-associated loci that have been reported in the descents of European, including \textit{FUT2}, \textit{TCN1}, \textit{MUT} and \textit{CUBN} (4–6). In our study, we clearly replicated the association of \textit{FUT2} loci and identified a novel SNP on \textit{FUT2} that is Asian-specific. In addition, the associations with VitB12 levels for the other three loci (\textit{TCN1}, \textit{MUT} and \textit{CUBN}) were also significant in our first stage of GWASs (\( P \)-value ranges from \( 7.1 \times 10^{-3} \) to \( 4.06 \times 10^{-4} \); Table 3) and with the same direction of association effect. Although the association did not reach a genome-wide significant level, it may due to the relative modest effect of these three loci in the Chinese population. Therefore, the previously reported VitB12-associated loci were significant in both Europeans and Asians. Similarly, the four novel loci identified from our Chinese population may not be race-specific, either. The novel SNPs might also be associated with VitB12 levels in Europeans, but with smaller effect sizes. However, this hypothesis needs to be tested in populations of European descent. We hope that the publication of our study will encourage more replication studies in European populations and, as well as other races, to test whether the novel loci identified from the Chinese population are also significantly associated with VitB12 level in other ethnicities.

One of the study limitations was the relatively high serum level of VitB12 in the participants of our study, compared with other studies (4–6). However, inter-assay coefficients of variation for the first stage and second stage of samples were 2.0 and 2.5\%, respectively. This indicates that the measurement error was well controlled according to the protocol. In addition, individuals treated with VitB12 or taken VitB12 supplements were excluded from our study. As we know, cereals, poultry, beef and fish are good sources of VitB12. Our study samples were collected from Fangchenggang, which is a coastal city in southern China. Fish represents the most commonly consumed food in the daily diary. Therefore, we speculated that the high consumption of fish as well as other foods that are in rich in VitB12 may potentially explain the relative high VitB12 level for the participants in our study.

In summary, we successfully identified four novel loci, including \textit{FUT6}, \textit{MS4A3}, \textit{5q32}, \textit{CLYBL} that were significantly associated with serum level of VitB12 in a Chinese population using a GWAS approach. We also identified one novel non-synonymous SNP in the previously reported \textit{FUT2} gene in Chinese population. The new loci identified offer new insights into the biochemical pathways involved in determining the serum level of VitB12 and provide opportunities to better delineate the role of VitB12 in health and disease.

**MATERIALS AND METHODS**

**Study participants**

The parent study population of the current study is from the Fangchenggang Area Male Health and Examination Survey (FAMHES) (23). Briefly, participants in FAMHES were recruited from Fangchenggang city, Guangxi, Southern China, with no prior history of cardiovascular disease, or other major chronic diseases. Every participant in the study would fulfill physical examinations in the Medical Centre of Fangchenggang First People’s Hospital from September 2009 to December 2009 and provide a blood sample at the same time. Initially about 4364 healthy men were asked to attend this study and 4303 individual (98.6\%) consented with age-span from 17- to 88-year-old.

The recruited subjects, with a total of 1999 members, in our stage 1 study from FAMHES cohort were limited to age 20–69 years old and were all self-reported southern Chinese Han ethnicity. The replication subjects in the stage 2 consisted of 1496 men age 20–70 years old and were recruited in conjunction with health examinations that were performed at three collaborating hospitals in Guangxi, China.

Written informed consents to participate in the present study were obtained from all of the subjects, and then standardized health questionnaires were used through a face-to-face interview conducted by trained physicians. Collected data included demographic, lifestyle characteristics (smoking, alcohol consumption), health status and family history and medical histories. The subjects who self-reported with diabetes mellitus, coronary heart disease, stroke, hyperthyroidism, rheumatoid arthritis and tumors were excluded from the study populations, moreover, and the detected samples with abnormal liver function. The study was approved by the Ethics and Human Subject Committee of Guangxi Medical University.

**Measurement of VitB12**

Overnight (\( \geq 8 \) h) fasting venous blood specimens were obtained and serum samples were extracted. Serum VitB12 was measured with electrochemiluminescence immunoassay on COBAS 6000 system E601 (Elecsys module) immunoassay analyzer (Roche Diagnostics, GmbH, Mannheim, Germany) with the same batch of reagents. Reference values of VitB12 are 197–866 pg/ml (for Europe) and 243–894 pg/ml (for USA). All assays and quality control (QC) were performed according to the manufacturer’s instructions and standard QC protocols. Specifically, Elecsys PreciControl Anemia (PCA) 1, 2 and 3, i.e. low, moderate and high concentrations, were used as internal known standards and were run before each batch of measurement. The measurements of the concentrations of the known internal standards were below two standard deviations for each batch. Inter-assay coefficients of variation were 2.0\% for the first stage and 2.5\% for the second stage.
SNP genotyping

Two different platforms were used for SNP genotyping. The Illumina Omini One platform was used for the genome-wide assay of samples in the first stage. For SNPs that were followed in the second stage, the iPLEX of Sequenom platform was used for genotyping (Sequenom, Inc., San Diego, CA, USA). Polymerase chain reaction and extension primers were designed using MassARRAY Assay Design 3.1 software (Sequenom, Inc.). Genotyping procedures were performed according to the manufacturer’s iPLEX Application Guide (Sequenom, Inc.). All genotyping reactions were performed in 384-well plates. Each plate included a duplicate for three or four subjects selected at random, as well as six to nine negative controls in which water was substituted for DNA. The average concordance rate was 99.8%.

Statistical analysis

QC procedures were first applied to 1999 individuals who were genotyped using the Illumina Omni-Express platform. A total of 1999 individuals passed the call rate of 95% and were used in the final statistical analysis. We then applied the following QC criteria to filter SNPs: $P < 0.001$ for the Hardy–Weinberg equilibrium test, minor allele frequency (MAF) < 0.01 and genotype call rate < 95%, and finally 709,211 SNPs retained. The IMPUTE computer program (24) was then used to infer the genotypes of SNPs (e.g. SNPs catalogued in Hapmap Phase II CHB population (24)) that remained significant at $10^{-5}$ after applying the IMPUTE software. A total of 1940,245 SNPs remained in our final analysis after applying the same QC criteria, as mentioned above.

The associations between serum VitB12 level and SNP genotypes were evaluated using a linear regression model assuming additive effects of the alleles (0, 1 and 2). In the regression models, the age, logfolic acid and BMI were adjusted as covariates for VitB12. The PLINK software package was used to perform this statistical analysis (25). The EIGENSTRAT software was applied for the evaluation of population stratification by a principal component approach (26). The top two eigenvectors were adjusted as covariates in the linear regression analysis. Q–Q plots were generated using R package. For regions with multiple SNPs that were significant at $P < 10^{-5}$, a multivariate linear regression analysis was applied to test the independence of the respective SNPs, and only SNPs that remained significant at $10^{-5}$ in the multivariate analysis were selected; thus yielding one SNP per region to be followed in the second stage. The combined analysis of two-stage data was performed using a linear regression, adjusting for the covariates and stage information.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

We thank Z.-J. Liang, D.-Y. Shi and Z. Huang from Fanchenggang First People’s Hospital, X.-H. Huang from Guigang People’s Hospital and M. Liu from Yulin First People’s Hospital for their contribution to the recruitment of study subjects. We thank X.-W. Zou, H.-C. Zheng and O. Li at the Genery Biotechnology (Shanghai) Co., Ltd, for their assistance in the genotyping. Finally, we thank all study subjects for participating in the study.

Conflict of Interest statement. None declared.

FUNDING

This work was partially supported by the General Program of National Natural Science Foundation of China (30945204, 30360124, 30260110), the Guangxi Provincial Department of Finance and Education (2009GJCJ150) and intramural funding from Fudan-VARI Center for genetic Epidemiology and intramural funding from Fudan University Institute of Urology.

REFERENCES


