nature genetics

Common variants of *FUT2* are associated with plasma vitamin B_{12} levels

Aditi Hazra[1,2], Peter Kraft[1], Jacob Selhub[3], Edward L Giovannucci[2,4], Gilles Thomas[5], Robert N Hoover[5], Stephen J Chanock[5] & David J Hunter[1,2,4–6]

We identified a strong association ($P = 5.36 \times 10^{-17}$) between rs492602 in *FUT2* and plasma vitamin B_{12} levels in a genome-wide scan ($n = 1{,}658$) and an independent replication sample ($n = 1{,}059$) from the Nurses' Health Study. Women homozygous for the rs492602[G] allele had higher B_{12} levels. This allele is in strong linkage disequilibrium with the *FUT2* nonsecretor variant encoding W143X, suggesting a plausible mechanism for altered B_{12} absorption and plasma levels.

Plasma level of vitamin B_{12} is a modifiable quantitative trait associated with certain diseases[1]. Vitamin B_{12} found in meat (such as liver, shellfish, fish, poultry and eggs) and milk products[2] is composed of corrin and cobalt rings and is necessary for the formation of red blood cells, DNA synthesis during cell division and maintenance of the myelin nerve sheath. Deficiency in vitamin B_{12}, clinically associated with pernicious anemia, cardiovascular disease, cancer and neurodegenerative disorders is often related to poor intestinal B_{12} absorption[2] rather than direct dietary deficiency (the recommended adult intake for vitamin B_{12} is 2.4 μg/day). In the Nurses' Health Study (NHS), we previously observed that women in the lowest quartile of plasma vitamin B_{12} levels had marginally worse cognitive performance (based on a global score averaging six cognitive tests) than women in the highest quartile of plasma vitamin B_{12} and that combined folate and vitamin B_{12} deficiency was associated with the lowest cognitive performance[3].

Epidemiologic studies provide evidence for the association of genes and metabolites in the B-vitamin–mediated plasma one-carbon metabolic pathway with chronic diseases[2–4]. Rare high-penetrance mutations in genes in this pathway affect the ability to digest, absorb[5] and utilize vitamin B_{12} (ref. 2). However, common genetic variants in candidate genes have not been consistently associated with plasma vitamin B_{12} levels. Therefore, we conducted a genome-wide association study (GWAS) to identify loci that influence plasma vitamin B_{12} levels in 1,658 women genotyped with the HumanHap500 as part of the Cancer Genetic Markers of Susceptibility project (CGEMS; **Supplementary Methods** online). Participants were of self-reported European ancestry[6]. Detailed methods have been previously reported, including quality control assessment of genotypes with sample completion and SNP call rates, concordance rate, deviation from Hardy-Weinberg proportions in control DNA and final sample selection for association analysis of the NHS CGEMS population[6].

We tested association between each of the 528,134 SNP markers that passed quality control filters and log-transformed plasma vitamin B_{12} using linear regression, adjusted for age, total methyl intake (defined as total folate, dietary methionine and alcohol intake) and assay batch. We did not find any evidence for systematic bias in the distribution of *P* values for analyses with and without further adjustment for residual population structure using the top four principal components of genetic variation (**Supplementary Fig. 1** online),

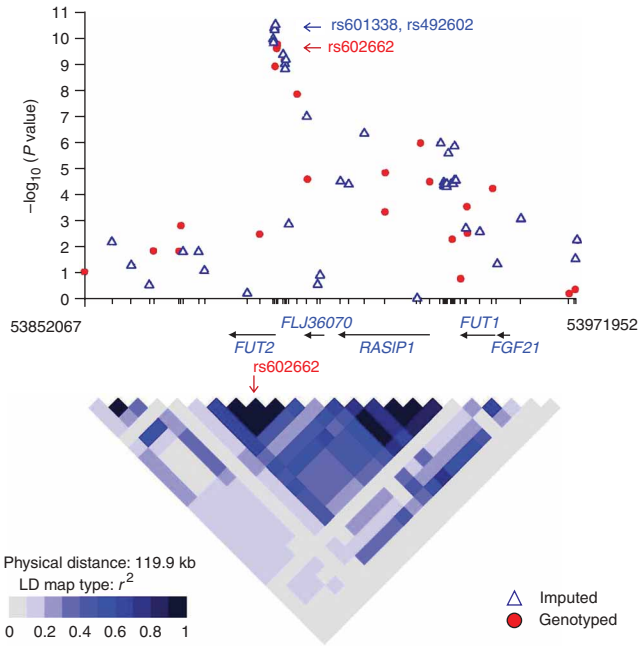

Figure 1 LD structure of chromosome 19. This panel shows the trend *P* values for association testing with plasma vitamin B_{12} from 24 genotyped (●) and 44 imputed (△) SNPs from the GWA study; SNPs on chromosome 19p13.3 are shown. LD plot is based on observed SNPs. All known genes are shown (National Center for Biotechnology Information build 36.1).

[1]Program in Molecular and Genetic Epidemiology, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, Massachusetts 02115, USA. [2]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 181 Longwood Avenue, Boston, Massachusetts 02115, USA. [3]Vitamin Metabolism and Aging Laboratory, Jean Mayer US Department of Agriculture Human Nutrition Research Center on Aging, Tufts University, Boston, Massachusetts 02111, USA. [4]Department of Nutrition, Harvard School of Public Health, 677 Huntington Avenue, Boston, Massachusetts 02115, USA. [5]Division of Cancer Epidemiology and Genetics, National Cancer Institute(NCI), US National Institutes of Health (NIH), Department of Health and Human Services (DHHS), Bethesda, Maryland 20892, USA. [6]Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA. Correspondence should be addressed to D.J.H. (dhunter@hsph.harvard.edu).

Received 18 March; accepted 26 June; published online 7 September 2008; doi:10.1038/ng.210

Table 1 Association of plasma vitamin B_{12} levels with *FUT2* genotypes in women participants from GWAS and replication studies

SNP association by study	Age	*N*	Allele freq.	Geometric mean (95% CI), pg/ml			Estimate (s.e.m.)[a]	*P* value[b]
rs602662			G (Gly) = 0.49	**Ser/Ser**	**Ser/Gly**	**Gly/Gly**		
GWAS								
NHS CGEMS	59	1,658		489.82 (472.24–508.06)	418.67 (407.74–429.90)	417.05 (401.29–433.43)	–0.08 (0.01)	6.54×10^{-10}
Replication								
Combined NHS CR adenoma and CR cancer case-control datasets	63	1,056		483.68 (461.78–506.63)	406.96 (393.69–420.67)	409.60 (390.22–429.93)	–0.08 (0.02)	1.13×10^{-06}
Total (GWAS and replication)				487.72 (473.80–502.06)	413.35 (404.83–422.06)	413.52 (401.14–426.27)	–0.08 (0.01)	3.52×10^{-15}
rs601338[c]			A (Trp) = 0.51[d]	**Nonsecretor (X/X)**	**Secretor (X/Trp)**	**Secretor (Trp/Trp)**		
GWAS								
NHS CGEMS	59	1,658		496.60 (477.89–516.04)	418.69 (407.76–429.92)	418.32 (403.36–433.84)	–0.08 (0.01)	4.11×10^{-10}
rs492602			A = 0.51[d]	**GG/GG**	**AG/AG**	**AA/AA**		
GWAS								
NHS CGEMS	59	1,637		496.02 (477.32–515.45)	419.95 (408.88–431.31)	416.97 (402.14–432.35)	–0.08 (0.01)	2.68×10^{-10}
Replication								
Combined NHS CR adenoma and CR cancer case-control datasets	63	1,059		491.13 (469.03–514.26)	407.67 (394.34–421.45)	406.61 (389.46–424.53)	–0.10 (0.02)	5.60×10^{-09}
Total (GWAS and replication)[e]				493.89 (479.42–508.81)	414.27 (405.69–423.04)	412.70 (401.33–424.40)	–0.09 (0.01)	5.36×10^{-17}

[a]Estimates (regression coefficients) calculated from linear regression adjusted for age using log-transformed plasma vitamin B_{12} (same estimates obtained when adjusting for age, assay batch and total dietary methyl status). [b]GWAS and replication *P* values are calculated from linear regression adjusted for age. [c]Imputed distribution for nonsense polymorphism, rs601338 (encoding W143X). [d]Major allele according to HapMap CEU data. [e]Dominant model *P* value for the joint analysis: rs602662 = 1.35×10^{-39} ; rs492602 = 8.26×10^{-45}. CR, colorectal.

compatible with no confounding of SNP-metabolite associations due to population stratification.

In the initial GWAS, the SNPs on chromosome 19p13.3 accounted for the excess of *P* values $<10^{-7}$ (**Supplementary Fig. 1**, quantile-quantile plot). The strongest association with plasma vitamin B_{12} was for rs602662 ($P_{trend} = 6.54 \times 10^{-10}$; **Fig. 1**, **Table 1** and **Supplementary Fig. 2** online), a nonsynonymous SNP in *FUT2* on chromosome 19p13.3, which has a minor allele frequency of 0.49 (**Table 1**; for HapMap frequencies, see **Supplementary Table 1** online). We independently replicated the association between plasma vitamin B_{12} and rs602662 using Taqman allelic discrimination assays in 1,059 women from two nested case-control studies drawn from the same cohort (**Table 1** and **Supplementary Table 2** online) with $P_{trend} = 1.13 \times 10^{-06}$. The joint analysis of the CGEMS scan data and the replication dataset supports the association of rs602662 with plasma vitamin B_{12} ($P_{trend} = 3.51 \times 10^{-15}$). We obtained similar results using untransformed plasma vitamin B_{12} values. A similar association was observed using the nonparametric Kruskal-Wallis test (joint analysis Kruskal-Wallis $P = 1.48 \times 10^{-21}$). Plasma vitamin B_{12} has an inverse correlation in this dataset with plasma homocysteine, an integrated marker of the one-carbon metabolism pathway (Spearman correlation coefficient –0.26, $P < 0.0001$). *FUT2* rs602662 also has a modest association with plasma homocysteine ($P_{trend} = 0.0085$ in the GWAS data and $P_{trend} = 0.0081$ in the replication data). The pattern of mean log-transformed plasma vitamin B_{12} levels by rs602662 genotype suggests a dominant genetic effect, with lower B_{12} levels among variant carriers (test comparing mean log-transformed vitamin B_{12} levels between variant carriers and noncarriers $P = 1.35 \times 10^{-39}$).

The SNP rs602662 is in strong linkage disequilibrium ($D' = 1$, $r^2 = 0.76$) with the *FUT2* nonsense SNP encoding W143X (rs601338; **Fig. 1**). We imputed the nonsense SNP in the initial GWAS samples using the observed chromosome 19 genotyping data augmented by data from the densely genotyped CEPH European HapMap samples and found that it was strongly associated with plasma vitamin B_{12} levels ($P_{trend} = 4.11 \times 10^{-10}$; **Fig. 1** and **Table 1**). As the nonsense SNP rs601338 could not be genotyped by Taqman because of a neighboring SNP (rs1800459) located one nucleotide upstream, we genotyped a proxy SNP, rs492602 (which is in perfect LD with rs601338 in HapMap data; **Supplementary Table 1** and **Supplementary Fig. 3**), in both the initial GWAS and the replication dataset. The association between rs492602 and plasma vitamin B_{12} was stronger than the association for rs602662 and plasma vitamin B_{12} (**Table 1**, rs492602 joint $P_{trend} = 5.36 \times 10^{-17}$; joint dominant model $P = 8.26 \times 10^{-45}$). The rs492602 SNP explains 2.5% of the variance in log-transformed plasma vitamin B_{12} levels using the log-additive genetic model and 3.5% of variance using the dominant genetic model.

Fucosylated carbohydrate structures are involved in a range of biological processes[7,8], including tissue development, angiogenesis,

fertilization, cell adhesion, inflammation and tumor metastasis. The classic human secretor locus *FUT2* encodes α,1,2-fucosyltransferase, which regulates expression of the Lewis ABO(H) histo-blood group antigens on the surface of epithelial cells and in body fluids and determines the secretion status of the ABO antigens[7]. Secretor status of this polymorphic protein was used by Mohr to provide the first autosomal linkage in humans between secretor factor and the Lutheran blood group; subsequently, secretor linkage was established with *APOE* and the locus associated with myotonic dystrophy[7]. The family of α,1,2-fucosyltransferases catalyze the addition of fucose in α,1,2-linkage to the galactose of type 1(Gal-β(1-3)-GlcNAc-R) and type 2 (Gal-β(1-4)-GlcNAc-R) disaccharide to form H type 1 and H type 2 antigens[7], respectively.

In the highly polymorphic *FUT2* gene[7], three SNPs, rs602662, rs492602 and rs601338 (W143X; nucleotide position 428), are in strong LD; we note that the variant encoding W143X is a nonsense mutation[7] and plausibly the causal variant for the association with plasma B_{12} levels. The variant encoding 143X is characteristic for the nonsecretor allele in Europeans and has an allele frequency of 0.46 in populations of European ancestry[7]. In Europeans, Africans and Iranians, the *FUT2* nonsense mutation encoding W143X is the primary nonsecretor allele, with a frequency of approximately 50%. In contrast, in Asian populations, the *FUT2* missense mutation encoding I129F (nucleotide position 385) is the primary nonsecretor allele[8]. In non-Asian populations, nonsecretors are frequently homozygous for the *FUT2* W143X polymorphism, resulting in an inactive *FUT2*. Individuals homozygous for the *FUT2* nonsecretor genotype seem to be resistant to infection with norovirus[9], suggesting that individuals homozygous for nonsecretor status may be unable to mediate host-microbe interactions[9].

Absorption of B_{12} requires the secretion of the glycoprotein intrinsic factor from the gastric cells, binding of intrinsic factor to vitamin B_{12} and a functional gastrointestinal absorption system[2]. The H-antigen synthesized by *FUT2*, Lewis ABO antigens and *FUT2* genotypes have all been reported to mediate *Helicobacter pylori* attachment to human gastric mucosa[10]. Atrophic gastritis is a consequence of *H. pylori* infection[11] and leads to reduced secretion of intrinsic factor[12,14]. The *FUT2* secretor status has been associated with both *H. pylori* infection and gastritis[10]; individuals with vitamin B_{12} malabsorption and low levels of serum vitamin B_{12} have higher seroprevalence of *H. pylori* infection[10]. These data suggest a potential mechanism by which vitamin B_{12} absorption[12–15] may be reduced in carriers of the secretor genotype owing to the sequelae of susceptibility to *H. pylori* infection compared to individuals with the nonsecretor genotype.

In summary, among generally well-nourished women, we found that common variation in the *FUT2* secretor gene is associated with plasma vitamin B_{12} at high levels of statistical confidence. Although the participants included in these analyses are a selected subset of women who had plasma vitamin B_{12} levels measured, numerous demographic and lifestyle factors are comparable between this sample and the overall NHS cohort. Insights gained from the study of plasma vitamin B_{12} are likely to have implications for the study of complex diseases such as cognitive decline, cancer and cardiovascular disease. Further study is required to investigate the biological basis of our reported association findings.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank H. Ranu, C. Chen and the staff at the Core Genotyping Facility at the National Cancer Institute for their expertise. This research is supported by the National Institutes of Health Research Grants U54 CA100971, P01 CA87969, P01 CA55075, U01 CA098233, R01 CA 065725 and CA070817. A.H. is supported in part by training grant NIH T-32 CA 09001-30.

AUTHOR CONTRIBUTIONS

This study is a joint effort of the Nurses' Health Study investigators (A.H., P.K., E.L.G. and D.J.H.) and Cancer Genetics Markers of Susceptibility study investigators (G.T., R.N.H., S.J.C. and D.J.H.). A.H. and D.J.H. designed the study; J.S. measured the plasma vitamin B_{12} and homocysteine levels; A.H. performed the analyses; P.K. supervised the statistical analyses; A.H. wrote the manuscript and prepared the tables, figures and supplementary material, with input from all co-authors.

Published online at http://www.nature.com/naturegenetics/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Siva, A. *et al. QJM* **100**, 495–499 (2007).
2. Watanabe, F. *Exp. Biol. Med. (Maywood)* **232**, 1266–1274 (2007).
3. Kang, J.H. *et al. Epidemiology* **17**, 650–657 (2006).
4. Dahlin, A.M. *et al. Int. J. Cancer* **122**, 2057–2061 (2008).
5. Tanner, S.M. *et al. Hum. Mutat.* **23**, 327–333 (2004).
6. Hunter, D.J. *et al. Nat. Genet.* **39**, 870–874 (2007).
7. Kelly, R.J. *et al. J. Biol. Chem.* **270**, 4640–4649 (1995).
8. Koda, Y. *et al. Genetics* **158**, 747–756 (2001).
9. Lindesmith, L.C. *et al. PLoS Med.* **5**, e31 (2008).
10. Carmel, R. *et al. J. Lab. Clin. Med.* **109**, 454–463 (1987).
11. Azevedo, M. *et al. J. Pathol.* **215**, 308–316 (2008).
12. van Oijen, M.G. *et al. J. Nutr. Sci. Vitaminol. (Tokyo)* **50**, 305–308 (2004).
13. Dholakia, K.R. *et al. World J. Gastroenterol.* **11**, 7078–7083 (2005).
14. Annibale, B. *et al. Dig. Liver Dis.* **34** (Suppl. 2), S72–S77 (2002).
15. Tamura, A. *et al. Am. J. Gastroenterol.* **97**, 861–866 (2002).