SHORT COMMUNICATION

Genome-wide association study of serum albumin:globulin ratio in Korean populations

Kyung-Won Hong¹, Hyun-Seok Jin², Daesub Song¹, Hye-kyoung Kwak¹, Sung Soo Kim¹ and Yeonjung Kim¹

Low albumin:globulin (A/G) ratios are associated with vascular adverse events, nephrotic syndrome and autoimmune disease. Genome-wide association studies (GWASs) have been identifying genetic variants associated with total serum protein, serum albumin and globulins, but A/G ratio has never been considered the target phenotype. To identify the genetic basis of the A/G ratio, we performed a GWAS on A/G ratio in 4205 individuals from the Ansan cohort and confirmed the results in 4637 subjects from the Ansong cohort. The single-nucleotide polymorphism (SNP) genotypes of Affymetrix SNP array 5.0 were obtained from the Korean Association Resource Consortium, and we selected 290 659 common SNPs with a minor allele frequency >0.05. Genetic factors for A/G ratio were analyzed by linear regression analysis, controlling for age, sex, body mass index, smoking status and alcohol drinking status as covariates. From the GWAS of the Ansan cohort, we identified two significant genome-wide signals (P -values < 5×10^{-8}) and 36 moderate signals (P -value < 1.0×10^{-4}). These 38 signals were tested in the Ansong population. Eleven SNPs from six loci (*GALNT2*, *IRF4*, *HLA-DBP1*, *SLC31A1*, *FADS1* and *TNFRSF13B*) were replicated, with P -values < 0.05. The most compelling association was observed in the *TNFRSF13B* locus on chromosome 17p11.2 (SNP: rs4561508), with an overall combined P -value = 7.80×10^{-24} . The other significant signal was observed on chromosome 11q12.2—the *FADS1* locus (SNP: rs174548)—with an overall combined P -value = 3.54×10^{-8} .

Journal of Human Genetics (2013) 58, 174–177; doi:10.1038/jhg.2012.130; published online 10 January 2013

Keywords: albumin:globulin ratio; *FADS1*; GWAS; *TNFRSF13B*; total protein

The albumin:globulin ratio (A/G ratio) is a biochemical parameter that is used to interpret changes in serum proteins that accompany disease.¹ The primary clinical manifestation of the A/G ratio is a reduction due to a decrease in serum albumin, and the subsequent increase in serum globulins.² This pattern is the classical change that is expected to accompany liver disease when serum albumin decreases below normal levels.³ A low A/G ratio is associated with all-cause mortality after non-ST (electrocardiogram S and T wave interval) elevation myocardial infarction,⁴ nephrotic syndrome⁵ and autoimmune disease.⁶

Recent genome-wide association studies (GWASs) have identified many loci that are linked to common diseases and quantitative traits.^{7,8} Previous GWASs have reported candidate genes for total serum proteins, serum albumin and globulins, but A/G ratio was not considered the target phenotype.^{9,10,11} Thus, we performed a GWAS using two population-based cohorts, comprising 8842 Korean individuals.

Study subjects were selected from an ongoing population-based cohort, as part of the Korean Genome and Epidemiology Study. Participants were recruited from residents in two cities (Ansong and Ansan) in Gyeonggi-do province, Korea. The Ansan cohort consisted

of 4637 individuals and was used as the discovery data set for GWAS. The Ansong cohort comprised 4205 persons and was the validation data set for the replication study. Written informed consent was obtained from all participants, and this research project was approved by the institutional review board of the Korea National Institute of Health.

Total serum protein and serum albumin were measured on an ADIVA1650 (Siemens, Tarrytown, NY, USA) and analyzed at a commercial analytical service center (SCL Co., Seoul, Korea). Globulin levels were calculated as the difference between total protein and albumin.

Genomic DNA was genotyped using the Affymetrix Genome-wide Human single-nucleotide polymorphism (SNP) array 5.0 at final concentrations of 100 g ml⁻¹ and 500 ng (Affymetrix, Inc., Santa Clara, CA, USA). The quality control steps have been described.¹² Briefly, SNPs with a missing genotype call rate >0.1 or a minor allele frequency <0.01 and in Hardy–Weinberg equilibrium ($P < 1 \times 10^{-6}$) were excluded; ultimately, we used 333 651 SNPs that were genotyped in the Korean Association Resource study.¹² For this study, we also removed SNPs on the X and Y chromosomes and those with minor allele frequency <5%; thus, 290 659 markers were used for the

¹Division of Epidemiology and Health Index, Center for Genome Science, Korea Centers for Disease Control and Prevention, Chungcheongbuk-do, Korea and ²Department of Medical Genetics, School of Medicine, Ajou University, Suwon, Korea

Correspondence: Dr Y Kim, Division of Epidemiology and Health Index, Center for Genome Science, Korea Center for Disease Control and Prevention, no. 64-3 Yeonje-ri, Gangneung-myeon, Gangneung-gun, Gangwon-do 363-951, Korea.

E-mail: kimye@korea.kr

Received 15 July 2012; revised 4 October 2012; accepted 18 October 2012; published online 10 January 2013

GWAS. For the fine mapping of the identified loci, we included the imputed SNPs in the signal plots. The detailed imputation procedure has previously been described in Hong *et al.*,¹³ which were previously used in the Korean Association Resource study. Briefly, the genotypes of the individuals were imputed using IMPUTE¹⁴ based on the International HapMap Phase II JPT (Japanese in Tokyo) + HCB (Han-Chinese in Beijing) panel.

The demographics and serum protein traits between Ansan and Ansong subjects were compared by χ^2 -test and *t*-test. The GWAS on the A/G ratio of the Ansan cohort was performed by linear regression, controlling for covariates, such as age, sex, body mass index, smoking status and alcohol drinking. SNPs that had a *P*-value $< 1 \times 10^{-4}$ were selected. The selected SNPs were validated in the Ansong cohort. Combined β - and *P*-values were obtained for the SNPs with *P*-values < 0.05 in Ansong cohort. Most association analyses were performed using PLINK (version 1.07; <http://pngu.mgh.harvard.edu/~purcell/plink/>) and SAS (version 9.1; SAS institute inc., Cary, NC, USA).

Table 1 describes the demographic features, body mass index, smoking status, alcohol drinking status and serum protein traits of the cohorts. The concentrations of all protein traits were higher in the Ansan cohort compared with the Ansong cohort. The mean age of the Ansan cohort was significantly lower than that of Ansong subjects. Thus, the difference in serum protein traits might be attributed to the older mean age in Ansong subjects, as reported by a Japanese epidemiological study, in which decreasing tendency of serum protein concentrations.¹⁵

Quantile–quantile plots and minus $\log_{10}(P)$ values were plotted against chromosomal position on Manhattan plots, as shown in Supplementary Figure 1 and Figure 1, respectively. The measured genomic inflation factor was 1.014. Thirty-eight SNPs in 18 loci had a *P*-value $< 10^{-4}$, two of which had genome-wide significance (*P*-value $< 5 \times 10^{-8}$). The results of the 38 SNPs can be found in Supplementary Table 1.

Subsequently, 38 SNPs were extracted from the genome-wide SNPs in the Ansong cohort and examined in the replication study. In all, 5 of the 38 SNPs were significantly replicated in Ansong subjects, with $P < 1.3 \times 10^{-3}$ (Bonferroni correction *P*-value criteria = 0.05/38), and 6 SNPs were weakly replicated with $P < 0.05$. The 11 SNPs lay in six loci (GALNT2, IRF4, HLA-DPB1, SLC31A1, FADS1 and TNFRSF13B), and the top SNPs in each locus are listed in Table 2. The signal plots for six replicate loci are shown in Supplementary Figures 2A–F.

Of the six loci, TNFRSF13B (rs4561508) and FADS1 (rs174548) showed genome-wide significance (overall $P < 5 \times 10^{-8}$) in the discovery cohort, and remained significant in the replication data

set. To compare the genetic effects on the other protein traits, we tested the six SNPs to the total serum protein, serum albumin and globulins in Table 2. The TNFRSF13B (rs4561508) and FADS1 (rs174548) revealed significant association with globulins, and the combined *P*-values were 3.36×10^{-21} and 1.37×10^{-8} , respectively. These results were similar to the association with A/G ratio.

In this study, we identified two significant loci and four suggestive loci that affect the A/G ratio. The most significant SNP resided in the TNFRSF13B gene, which was reported in two Japanese GWAS studies as the significant loci with regard to total serum protein concentration⁹ or globulins.¹¹ The SNPs rs4273077 and rs4985726 lie ~ 500 bp and ~ 5 kbp upstream, respectively, of rs4561508—identified in the current study. The linkage disequilibrium (LD) was moderate between rs4985726 and rs4561508 ($r^2 = 0.608$), and perfect LD between rs4985726 and 4561508 ($r^2 = 0.1000$).

The TNFRSF13B gene encodes a lymphocyte-specific member of the tumor necrosis factor receptor superfamily. It interacts with

Figure 1 Mahattan plot for the GWAS of A/G ratio. SNPs were plotted based on their physical chromosomal positions (horizontal axis) together with their $-\log_{10}(P)$ -values in the GWAS (vertical axis). The dotted horizontal line is the suggestive threshold of $P = 1 \times 10^{-4}$, and the solid horizontal line is the genome-wide significance threshold of $P = 5 \times 10^{-8}$.

Table 1 Basic characteristics of biochemical traits in the Ansan and Ansong cohorts

	Ansan	Ansong	P-value
Sample size	4637	4205	
Gender (males/females)	2374/2263	1809/2396	< 0.05
Age (year (range))	49.08 (39–69)	55.67 (39–70)	< 0.05
Total protein (mean $\pm$ s.d. g dl ⁻¹)	7.43 $\pm$ 0.45	7.18 $\pm$ 0.42	< 0.05
Serum albumin (mean $\pm$ s.d. g dl ⁻¹)	4.34 $\pm$ 0.35	4.15 $\pm$ 0.27	< 0.05
Serum globulin (mean $\pm$ s.d. g dl ⁻¹)	3.09 $\pm$ 0.32	3.04 $\pm$ 0.35	< 0.05
Albumin/globulin (mean $\pm$ s.d.)	1.42 $\pm$ 0.20	1.38 $\pm$ 0.19	< 0.05
Body mass index (mean $\pm$ s.d. kg m ⁻²)	24.5 $\pm$ 3.29	24.7 $\pm$ 2.96	< 0.05
Smoking state (never/quit smoking/current smoking)	2492/512/1105	2651/842/1123	< 0.05
Alcohol drinking (never/quit drinking/current drinking)	2068/320/1749	1995/243/2390	< 0.05

P-values were obtained by χ^2 -test for gender and Student's *t*-test for quantitative variables.

Table 2 Repeat associated SNPs in the Ansan and Ansung cohorts and the further association results with total proteins, albumin and globulins

CHR	SNP	Gene	BP	A1	A2	MAF	Phenotypes	Discovery GWAS (Ansan), n = 4637			Replication study (Ansung), n = 4205			Combined analysis, n = 8842		
								β	s.e.	P	β	s.e.	P	β	s.e.	P
1	rs3761944	GALNT2	228455612	C	G	0.224	A/G ratio	0.020	0.005	6.21×10^{-5}	0.012	0.005	0.019	0.016	0.004	7.50×10^{-6}
							Total protein	-0.006	0.011	0.590	-0.005	0.11	0.687	-0.006	0.008	0.485
							Serum albumin	0.22	0.008	9.00×10^{-3}	0.017	0.007	0.023	0.019	0.006	8.31×10^{-4}
							Globulins	-0.28	0.008	5.72×10^{-4}	-0.021	0.009	0.025	-0.025	0.006	6.60×10^{-5}
6	rs6927090	IRF4	197145	A	C	0.282	A/G ratio	-0.017	0.004	7.52×10^{-5}	-0.012	0.005	0.011	0.015	0.003	3.95×10^{-6}
							Total protein	0.022	0.01	0.030	0.006	0.01	0.532	0.014	0.007	0.046
							Serum albumin	-0.008	0.007	0.297	-0.008	0.007	0.237	-0.008	0.005	0.125
							Globulins	0.03	0.007	4.40×10^{-5}	0.014	0.008	0.092	0.022	0.006	5.26×10^{-5}
6	rs3117229	HLA-DBP1	33164047	G	A	0.493	A/G ratio	0.015	0.004	8.57×10^{-5}	0.009	0.004	0.027	0.012	0.003	1.44×10^{-5}
							Total protein	-0.005	0.009	0.611	-0.017	0.009	0.059	-0.011	0.006	0.101
							Serum albumin	0.017	0.007	0.010	0.003	0.006	0.566	0.011	0.005	0.021
							Globulins	-0.022	0.006	0.001	-0.021	0.008	0.006	-0.021	0.005	1.81×10^{-5}
9	rs6477998	SLC31A1	115037070	C	A	0.467	A/G ratio	0.016	0.004	9.64×10^{-5}	0.01	0.004	0.021	0.013	0.003	7.66×10^{-6}
							Total protein	0.004	0.009	0.699	-0.009	0.009	0.319	-0.002	0.007	0.779
							Serum albumin	0.021	0.007	0.003	0.009	0.006	0.114	0.016	0.005	5.77×10^{-4}
							Globulins	-0.017	0.007	0.009	-0.019	0.008	0.014	-0.018	0.005	3.47×10^{-4}
11	rs174548	FADS1	61327924	G	C	0.319	A/G ratio	-0.017	0.004	3.37×10^{-5}	-0.015	0.004	7.76×10^{-4}	0.017	0.003	3.54×10^{-8}
							Total protein	0.017	0.009	0.078	0.028	0.01	0.004	0.02	0.007	0.003
							Serum albumin	-0.01	0.007	0.154	-0.005	0.006	0.433	-0.009	0.005	0.055
							Globulins	0.027	0.007	8.92×10^{-5}	0.033	0.008	4.83×10^{-5}	0.03	0.005	1.37×10^{-8}
17	rs4561508	TNFRSF13B	16789475	T	C	0.335	A/G ratio	-0.03	0.004	6.09×10^{-13}	-0.032	0.004	2.67×10^{-13}	0.031	0.003	7.80×10^{-24}
							Total protein	0.033	0.009	3.97×10^{-4}	0.03	0.01	0.002	0.034	0.007	1.01×10^{-6}
							Serum albumin	-0.018	0.007	0.011	-0.017	0.006	0.007	-0.016	0.005	0.001
							Globulins	0.051	0.007	4.73×10^{-14}	0.047	0.008	6.25×10^{-9}	0.05	0.005	3.36×10^{-21}

Abbreviations: A1, minor allele and coded allele; A2, major allele; BP, base pair based on human genome build no. 36; CHR, chromosome; Gene, nearby gene; GWAS, genome-wide association study; MAF, minor allele frequency; Replicated SNP, the number of SNPs in each locus; SNP, single-nucleotide polymorphism (rsID of dbSNP).

calcium modulator and cyclophilin ligand, activates the transcription factors NFAT, AP1 and NF-kappa-B, and regulates humoral immunity by interacting with a tumor necrosis factor ligand. TNFRSF13B mutations have been reported to be associated with common variable immunodeficiency¹⁶ and hypogammaglobulinemia.¹⁷ Among the suggestive loci, IRF4 and HLA-DBP1 are important inflammation factors. Thus, our study suggests that A/G ratio predicts genetic predisposition to susceptibility for liver inflammation.

The FADS1 (FAD synthetase 1) gene was first identified in this study as a candidate gene for the A/G ratio. FAD synthetase is a critical member of the electron transport chain and is believed to mediate the cellular oxidative phosphorylation pathway.^{18,19} This gene was identified as the antigenic target of immunoglobulin A.²⁰ The rs174548 SNP in *FADS1* has been reported to have robust associations with high-density lipoprotein cholesterol and triglyceride levels in European studies.^{21,22} Cholesterol in the blood is synthesized in the liver and is a factor of cerebral infarction, myocardial infarction and liver function. Thus, our association results for rs174548 and A/G ratio are relevant to the previous two GWAS studies.

Serum protein measurements can be expected to provide greater detail on the potentially affected disease pathway. Thus, we hope that our findings increase our understanding of clinical end points, such as coronary artery disease and kidney disease.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

- 1 Al-Joudi, F. S. & Wahab, N. A. The utilization of an index for serum globulin compensation in diseases associated with decreased serum albumin. *Med. J. Malaysia* **59**, 495–501 (2004).
- 2 Henry, J. B. *Clinical Diagnosis and Management by Laboratory Methods*. (W B Saunders Co., Philadelphia, 1996).
- 3 Dufour, D. R., Lott, J. A., Nolte, F. S., Gretch, D. R., Koff, R. S. & Seeff, L. B. Diagnosis and monitoring of hepatic injury. II. Recommendations for use of laboratory tests in screening, diagnosis, and monitoring. *Clin. Chem.* **46**, 2050–2068 (2000).
- 4 Azab, B., Bibawy, J., Harris, K., Khoeiry, G., Akeman, M., Selim, J. *et al.* Value of albumin-globulin ratio as a predictor of all-cause mortality after non-ST elevation myocardial infarction. *Angiology* (in press) (2012).

- 5 Giles, H. M., Pugh, R. C. B., Darmady, E. M., Stranack, F. & Woolf, L. I. The nephrotic syndrome in early infancy: a report of three cases. *Arch. Dis. Child.* **32**, 167–180 (1957).
- 6 Al-Joudi, F. S. Prognostic value of an index for serum globulin compensation in colon and breast cancers. *Singapore Med. J.* **46**, 710–713 (2005).
- 7 Ku, C. S., Loy, E. Y., Pawitan, Y. & Cha, K. S. The pursuit of genomewide association studies: where are we now? *J. Hum. Genet.* **55**, 195–206 (2010).
- 8 Visscher, P. M., Brown, M. A., McCarthy, M. I. & Yang, J. Five years of GWAS discovery. *Am. J. Hum. Genet.* **90**, 7–24 (2012).
- 9 Kim, Y. J., Go, M. J., Hu, C., Hong, C. B., Kim, Y. K., Lee, J. Y. *et al.* Large-scale genomewide association studies in east Asians identify new genetic loci influencing metabolic traits. *Nat. Genet.* **43**, 990–995 (2011).
- 10 Kamatani, Y., Matsuda, K., Okada, Y., Kubo, M., Hosono, N., Daigo, Y. *et al.* Genomewide association study of hematological and biochemical traits in a Japanese population. *Nat. Genet.* **42**, 210–215 (2010).
- 11 Osman, W., Okada, Y., Kamatani, Y., Kubo, M., Matsuda, K. & Nakamura, Y. Association of common variants in TNFRSF13B, TNFSF13, and ANXA3 with serum levels of non-albumin protein and immunoglobulin isotypes in Japanese. *PLOS one* **7**, e32683 (2012).
- 12 Cho, Y. S., Go, M. J., Kim, Y. J., Heo, J. Y., Oh, J. H., Ban, H. J. *et al.* A large-scale genomewide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat. Genet.* **41**, 527–534 (2009).
- 13 Hong, K. W., Lim, J. E., Kim, Y. J., Cho, N. H., Shin, C. & Oh, B. K. A. R. E. Genomewide association study of blood pressure using imputed SNPs. *Genomics Inform.* **8**, 103–107 (2010).
- 14 Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
- 15 Chan, Y. C., Suzuki, M. & Yamamoto, S. A comparison of anthropometry, biochemical variables and plasma amino acids among centenarians, elderly and young subjects. *J. Am. Coll. Nutr.* **18**, 358–365 (1999).
- 16 Castigli, E., Wilson, S. A., Garibyan, L., Rachid, R., Bonilla, F., Schneider, L. *et al.* TACI is mutant in common variable immunodeficiency and IgA deficiency. *Nat. Genet.* **37**, 829–834 (2005).
- 17 Salzer, U., Chapel, H. M., Webster, A. D., Pan-Hammarstrom, Q., Schmitt-Graeff, A., Schlesier, M. *et al.* Mutations in TNFRSF13B encoding TACI are associated with common variable immunodeficiency in humans. *Nat. Genet.* **37**, 820–828 (2005).
- 18 Barile, M., Passarella, S., Bertoldi, A. & Quagliariello, E. Flavin adenine dinucleotide synthesis in isolated rat liver mitochondria caused by imported flavin mononucleotide. *Arch. Biochem. Biophys.* **350**, 442–447 (1993).
- 19 Deluca, C. & Kaplan, N. O. Flavin adenine dinucleotide synthesis in animal tissues. *Biochem. Biophys. Acta* **30**, 6–11 (1958).
- 20 Lin, J., Diamanduros, A., Chowdhury, S. A., Scelsa, S., Latov, N. & Sadiq, S. A. Specific electron transport chain abnormalities in amyotrophic lateral sclerosis. *J. Neurol.* **256**, 774–782 (2009).
- 21 Gieger, C., Geistlinger, L., Altmaier, E., Hrabé de Angelis, M., Kronenberg, F. *et al.* Genetics meets metabolomics: a genomewide association study of metabolite profiles in human serum. *PLoS Genet.* **4**, e1000282 (2008).
- 22 Waterworth, D. M., Ricketts, S. L., Song, K., Chen, L., Zhao, J. H., Ripatti, S. *et al.* Genetic variants influencing circulating lipid levels and risk of coronary artery disease. *Arterioscler. Thromb. Vasc. Biol.* **30**, 2264–2276 (2010).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)