

Postprint

Genetic determinants of circulating haptoglobin concentration

Nabila Kazmi¹, Yoshiro Koda², Ndeye Coumba Ndiaye³, Sophie Visvikis-Siest³, Matt Morton⁴,
Tom Gaunt¹, Ian Galea⁴

¹MRC Integrative Epidemiology Unit (IEU), Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

²Department of Forensic Medicine, Kurume University School of Medicine, Kurume 830-0011, Japan

³Université de Lorraine, Inserm, IGE-PCV, Nancy, France

⁴Clinical Neurosciences, Clinical & Experimental Sciences, Faculty of Medicine, University of Southampton, UK

Word count: 1512

Abstract

Haptoglobin (Hp) is a major plasma acute-phase glycoprotein, which binds free haemoglobin to neutralize its toxicity. The *HP* gene exists as two copy number variants (CNV), *HP1* and *HP2*, which differ in two ways: serum Hp level and functional differences in Hp protein products. Both mechanisms may underlie the *HP* CNV's influence on susceptibility and/or outcome in several diseases. A single nucleotide polymorphism rs2000999 has also been associated with serum Hp level. In a meta-analysis of three studies from England, France and Japan, we show that rs2000999's effect on circulating Hp level is independent from that of the *HP* CNV. The combined use of rs2000999 and the *HP* CNV can be an important genetic epidemiological tool to discriminate between the two potential mechanisms underlying differences between *HP1* and *HP2* alleles.

Keywords: ALSPAC, haptoglobin, copy number variant

Introduction

Haptoglobin (Hp) is a major plasma acute-phase glycoprotein. The main biological function of Hp is to bind free hemoglobin (Hb) to prevent the loss of iron and subsequent kidney damage following intravascular hemolysis¹; it may also have immunomodulatory properties¹. In humans, the *HP* locus is polymorphic. The basic structure of Hp is of an alpha and beta chain that form a monomer; a disulphide bond between the alpha chains of two monomers supports the formation of a dimer in its mature form. An intragenic duplication of exons 3 and 4 of the ancestral *HP* gene produced the *HP2* allele, after which recurrent deletions produced the modern *HP1* allele². Therefore individuals could be one of three genotypes: *HP1-1*, *HP2-1* or *HP2-2*³. This copy number variant (CNV) codes for an extra cysteine residue in the alpha chain that forms an additional disulphide bond; therefore trimeric, tetrameric and higher order polymers are formed in individuals carrying the *HP2* allele³.

The *HP* CNV may influence the susceptibility and/or outcome in several diseases, for example, diabetes mellitus^{4,5}, atherosclerosis⁶ and cardiovascular disease⁷, malignancies⁸ and infections⁹, with *HP2* conferring higher risk. The *HP2* allele is associated with a dose-dependent reduction in serum Hp level, compared to *HP1* (serum Hp level: *HP1-1* > *HP2-1* > *HP2-2*)¹⁰. The *HP1* and *HP2* alleles differ in their ability to clear Hb and may have other functional differences relating to protection from Hb-induced oxidative stress¹¹. Hence the differential effects of *HP1* and *HP2* alleles could be mediated by a difference in levels of circulating Hp, or functional differences in their protein products, or both.

A single nucleotide polymorphism (SNP) rs2000999 has also been associated with serum Hp level, resulting in a dose-dependent variation across the genotypes: GG > GA > AA¹². Since rs2000999 is downstream of the *HP* gene, its effect is probably mediated via linkage

disequilibrium (LD) with the upstream single base-pair deletion rs35283911¹⁰. Up to 45% of the variance in circulating Hp level is explained by rs2000999¹². It is likely that some of the effects of rs2000999 and the *HP* CNV on Hp level are independent of each other^{2,10,13}.

The aim of this study was to perform a meta-analysis to confirm the independence of rs2000999 from the *HP* CNV with respect to circulating Hp level. If rs2000999 truly has an independent effect on Hp level, the combined use of rs2000999 and the *HP* CNV can be an important genetic epidemiological tool to discriminate between the two potential mechanisms underlying differences between *HP1* and *HP2* alleles: Hp protein function or Hp protein expression level.

Methods

Participants

We included studies which: (1) genotyped both the *HP* CNV and rs2000999 and (2) genotyped the CNV using gold-standard PCR approaches (Table 1), rather than imputation from genome-wide array screening (GWAS). Three studies were identified at the time of study planning, and data was made available by the authors. The Avon Longitudinal Study of Parents and Children (ALSPAC) is a general population cohort study^{14,15}; circulating Hp level and *HP* genetic data was available from 325 participants (159 females and 166 males, Caucasian, age 18 ± 0.33 years). The ALSPAC study website (<http://www.bristol.ac.uk/alspac/researchers/our-data/>) contains details of all the data that is available through a fully searchable data dictionary and variable search tool. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. The STANISLAS Family Study (SFS) provided data from 500 French community-dwelling adult women of European ancestry (mean age 39.6 ± 4.1 years)^{12,16}. Soejima *et al* consisted of 385 Japanese patients scheduled

for blood transfusion (195 females and 190 males, Asian, adults of undetermined age)¹³. Table 1 lists the study characteristics including the methods used for *HP* genotyping and circulating Hp level assay.

Analysis

The contribution of the *HP* CNV and rs2000999 to circulating Hp levels was studied within each cohort using multivariate regression modelling in R (Hp level as outcome, *HP* CNV as exposure and rs2000999 as covariate). The interaction terms *HP* CNV * gender and rs2000999 * gender were considered, except in the female-only French cohort. The results of each cohort were combined in a fixed-effects meta-analysis conducted using the R package *metaphor*, with the inverse variance for weighting and the restricted maximum likelihood estimator method as heterogeneity estimator, and assuming a standard normal distribution to calculate confidence intervals (CI).

Results

Contribution of *HP* polymorphism to expression

Data from the two mixed gender British and Japanese cohorts and the female-only French cohort was used to assess the contribution of the two *HP* genetic polymorphisms to circulating Hp level. All cohorts were in Hardy-Weinberg equilibrium (HWE) for both genetic variants (Supplementary File 1, Table S1). We found evidence of LD between the two polymorphisms in each cohort (Supplementary File 1, Table S1).

To determine the contribution of *HP* genetic variation to circulating Hp level we performed multivariate linear regression where *HP* CNV was the exposure and rs2000999 was the

covariate and combined the effects of each cohort in a fixed-effects meta-analysis. The analysis identified that the rs2000999 minor allele (A) was associated with Hp change of -0.14 g/L (95% CI -0.20 to -0.09, $p=2.6\times10^{-7}$), and the *HP* CNV was also associated with Hp level (for each *HP2* allele, $\beta = -0.189$ g/L, 95% CI -0.24 to -0.14, $p=1.68\times10^{-15}$), Figure 1. There was no statistical evidence of between study heterogeneity (Cochrane Q p-values were > 0.4). The magnitude of effect was larger in univariate compared to multivariate analysis (cf Figures 1 and 2). The French and British studies, which restricted participants to healthy individuals only, broadly revealed the same results (Figures 1 and 2).

The main effect of gender was assessed in British and Japanese cohorts (British: rs2000999 $p = 0.69$, *HP* CNV $p = 0.08$; Japanese: rs2000999 $p = 0.10$, *HP* CNV $p = 0.96$). There was no interaction of gender with either the *HP* CNV or rs2000999 in determining Hp level (for meta-analysis: rs2000999 $p = 0.4$ and *HP* CNV $p = 0.1$, Figure 2).

The mean Hp level was 0.99 g/L (standard deviation (SD): 0.47) in the British cohort, 1.42 g/L (SD: 1.06) in the Japanese cohort and 1.06 g/L (SD: 0.43) in the French cohort. *HP2* and rs2000999:A were more frequent in Japanese *versus* Caucasian populations ($\chi^2 = 27.04$ and 71.80 respectively, degrees of freedom = 2 and 2 respectively, $p < 0.00001$ in both cases). In a multivariate linear regression, the effect of ancestry (Asian/European) on Hp level persisted despite controlling for the *HP* CNV, rs2000999 and gender (Supplementary File1, Table S2).

Discussion

This meta-analysis of *HP* genetic variation and circulating Hp concentration across British, Japanese and French populations showed that: (1) both rs2000999 and *HP* CNV contribute to

Hp level, (2) the effects of rs2000999 and the *HP* CNV are independent of each other, i.e. LD does not fully account for their combined effects on Hp level. The beta coefficients of univariate regression analyses were consistently higher than the respective multivariate analyses indicating that LD accounts for some of the effect seen in univariate analyses, but clear and strong independent effects were observed.

Other factors potentially influencing circulating Hp level include age, gender and ancestry. We could not include age as a covariate since participants in the British study were roughly of the same age, and the Japanese study did not collect age data. However, in the French cohort, age above 18 years did not influence Hp level^{12,17}. Gender may affect circulating Hp level¹⁸, so it was included as a covariate. Ancestry also impacts on circulating Hp level¹³. Some of the effect of ancestry could be due to inter-population differences in the frequencies of the *HP* CNV and rs2000999, but we show here that the effect of Asian/European status on Hp level persisted despite controlling for the *HP* CNV, rs2000999 and gender. This residual effect of ancestry could be due to other genetic determinants. However, comparison of studies performed in different populations needs to be interpreted with caution, since differences in Hp level could also be due to methodological differences in Hp assay technique between studies.

Strengths of this study include the large sample size and the multivariate nature of the analysis, controlling for gender, ancestry, *HP* CNV and *HP* SNP. Two of the populations (French and British) were unselected individuals allowing us to perform sensitivity analysis without the confounding factor of disease that exists in the Japanese study. Also the *HP* CNV was genotyped using a gold-standard PCR approach, not determined by imputation.

This study has a number of limitations. Other covariates potentially affecting Hp level, such as body mass index and cigarette consumption¹⁷, were not available. The Japanese cohort was not community-based since it was a population of patients requiring blood transfusion. This could have been responsible for the observed higher Japanese frequency of *HP2* and rs2000999:A, despite the population being in HWE. It is still possible that some of the higher Hp level in the Japanese cohort, unexplained by the *HP* CNV and rs2000999, was driven by underlying medical conditions in people requiring blood transfusion, as well as other population-specific genetic determinants.

In conclusion, we show that the *HP* CNV and SNP both contribute to circulating Hp level, independent of each other. Using these two *HP* genetic variations together is a valuable genetic epidemiological tool to dissect the mechanism underlying the differential effect of *HP1* and *HP2* alleles in disease, namely whether this is through Hp level or differences in function between *HP* polymers.

Contributors

Concept: TG, IG. Data contributors: NK, YK, NCN, SVS, TG. Analysis: NK, MM, IG. Manuscript: all authors.

Declaration of interests

Nothing to declare.

Funding statement

The UK Medical Research Council (MRC) and Wellcome Trust (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and TG, IG and NK are guarantors. A comprehensive list of grant funding is available on the ALSPAC website (<http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>). This research was funded by UK MRC grant MR/L01453X/1 (MM, IG) and by Cancer Research UK program grant C18281/A19169 (NK). Tom Gaunt receives funding from the UK MRC (MRC Integrative Epidemiology Unit, MC_UU_00011/4). ALSPAC GWAS data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe.

Acknowledgements

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

Figure 1: Effect of *HP*-CNV and rs2000999 on circulating Hp level in multivariate analyses. CNV was included in the model as a dependent variable and rs2000999 as a covariate whereas gender was included as an interaction term. The labels indicate the comparison being made; for instance CNV:HP2 indicates the change in Hp level in g/L for every HP2 allele (i.e. the beta coefficient). Gender:M = the gender is male.

Analyses

Odds Ratio [95% CI]

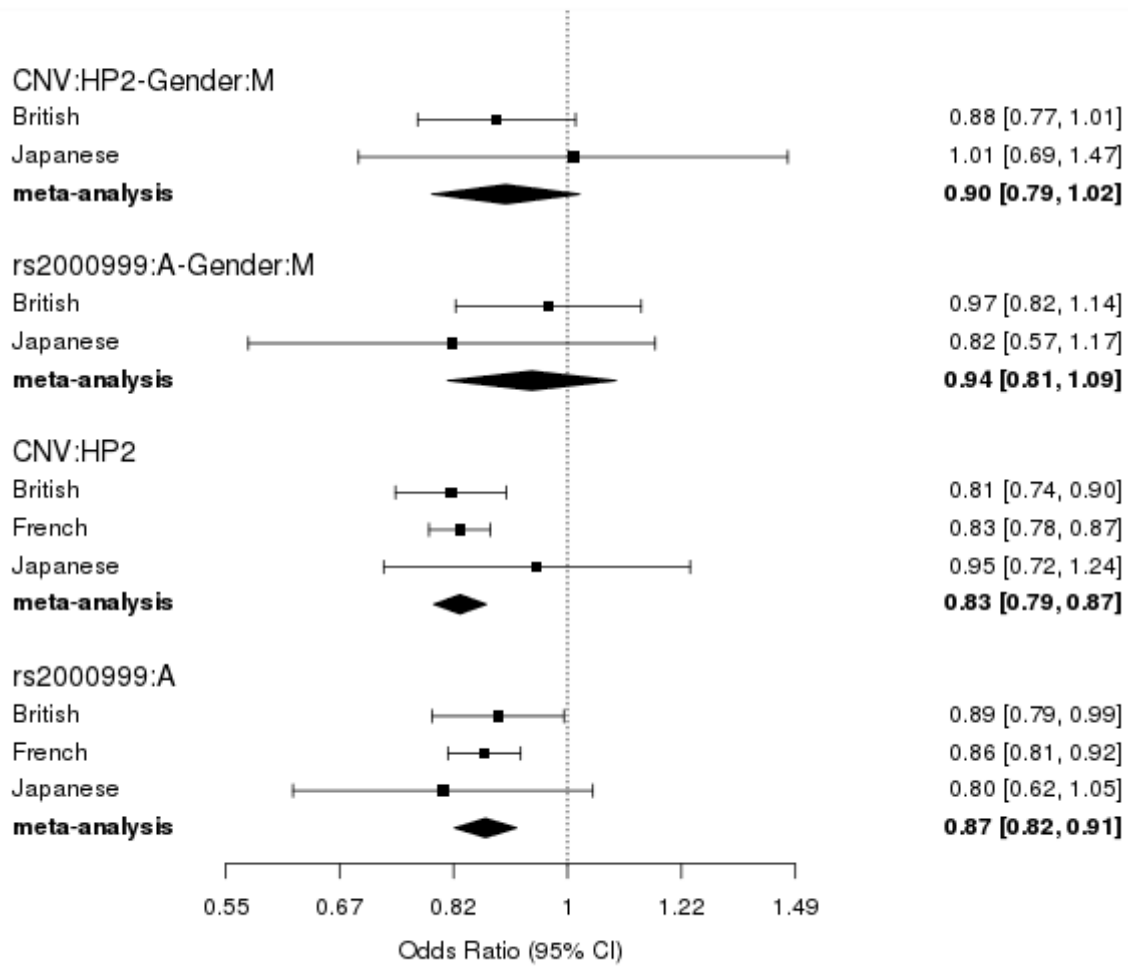






Figure 2: Effect of *HP*-CNV and rs2000999 on circulating Hp level in univariate analyses.





Analyses

Odds Ratio [95% CI]

CNV:HP2

British		0.75 [0.70, 0.80]
French		0.79 [0.75, 0.84]
Japanese		0.83 [0.70, 0.99]
meta-analysis		0.78 [0.75, 0.81]

rs2000999:A

British		0.79 [0.73, 0.86]
French		0.80 [0.75, 0.85]
Japanese		0.71 [0.60, 0.83]
meta-analysis		0.79 [0.75, 0.83]

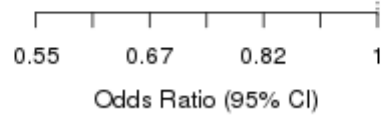


Table 1. Study populations

Study	N	Gender (% female)	Mean age (\pm SD)	Country	Source	HP CNV	rs2000999	Hp assay	Mean Hp level (\pm SD)	Ref
ALSPAC	325	49%	18 \pm 0.33	England	Community	ARCS ^b	GWSA	Immuno- turbimetry on plasma ^e	0.99 \pm 0.47	Fraser et al ^{14,15}
SFS	500	100%	39.6 \pm 4.1	France	Community	PCR ^c	GWSA	Immuno- nephelometry on plasma ¹²	1.07 \pm 0.43	Visvikis-Siest S et al ¹⁶
Japan	385	51%	Adult ^a	Japan	Hospital	qPCR ^d	qPCR	ELISA on serum ¹³	1.42 \pm 1.07	Soejima <i>et al</i> ¹³

Footnote: ARCS = amplification ratio control system, qPCR = quantitative polymerase chain reaction, GWSA = genome-wide SNP array,

ELISA = enzyme-linked immunosorbent assay. ^aAdult, exact age undetermined. ^bdoi:10.1093/nar/gkr046.

^c<http://clinchem.aaccjnls.org/content/48/9/1377>. ^ddoi:10.1373/clinchem.2008.113126.

^e<https://usdiagnostics.roche.com/products/03005593322/PARAM182/overlay.html>

Supplementary Table S1

	British	French	Japanese
rs2000999 - A/A	16	22	48
rs2000999 - A/G	107	176	197
rs2000999 - G/G	202	302	140
CNV - Hp1/Hp1	45	73	27
CNV - Hp1/Hp2	146	234	154
CNV - Hp2/Hp2	134	181	204
LD - D	-0.0612	-0.0694	-0.1007
LD - D'	0.788	0.811	0.9796
LD - Correlation coefficient	-0.31051	-0.344	-0.467
LD - Chi-square	62.669	115.502	167.976
LD-pvalue	2.44E-15	<0.001	<0.001
N	325	488	385
HWE-pvalue - rs2000999	0.7411	0.6955	0.1048
HWE-pvalue - CNV	0.6315	0.9242	0.8971

CNV = copy number variant

LD = linkage disequilibrium

HWE = Hardy–Weinberg equilibrium

Supplementary Table S2

Results of linear regression analysis where Hp serum level was the outcome and Race, CNV, rs2000999 and Gender were the exposures

	Coef	SE	pVal	CI
(Intercept)	3.782	0.0447	< 2.2e-16	3.694to3.87
Race-Japanese	1.6166	0.0456	< 2.2e-16	1.527to1.706
CNV	0.8441	0.0317	1.07E-07	0.782to0.906
rs2000999	0.8163	0.0349	7.60E-09	0.748to0.885
Gender	1.0133	0.0453	0.7697	0.925to1.102

References

1. Andersen CBF, Stodkilde K, Saederup KL, et al. Haptoglobin. *Antioxidants & redox signaling* 2017; **26**(14): 814-31.
2. Boettger LM, Salem RM, Handsaker RE, et al. Recurring exon deletions in the HP (haptoglobin) gene contribute to lower blood cholesterol levels. *Nature genetics* 2016; **48**(4): 359-66.
3. Maeda N, Yang F, Barnett DR, Bowman BH, Smithies O. Duplication within the haptoglobin Hp2 gene. *Nature* 1984; **309**(5964): 131-5.
4. Levy AP, Roguin A, Hochberg I, et al. Haptoglobin phenotype and vascular complications in patients with diabetes. *The New England journal of medicine* 2000; **343**(13): 969-70.
5. Bessa SS, Hamdy SM, Ali EM. Haptoglobin gene polymorphism in type 2 diabetic patients with and without nephropathy: An Egyptian study. *European journal of internal medicine* 2007; **18**(6): 489-95.
6. Shor M, Boaz M, Gavish D, Wainshtein J, Matas Z, Shargorodsky M. Relation of haptoglobin phenotype to early vascular changes in patients with diabetes mellitus. *The American journal of cardiology* 2007; **100**(12): 1767-70.
7. Levy AP, Hochberg I, Jablonski K, et al. Haptoglobin phenotype is an independent risk factor for cardiovascular disease in individuals with diabetes: The Strong Heart Study. *Journal of the American College of Cardiology* 2002; **40**(11): 1984-90.
8. Awadallah SM, Atoum MF. Haptoglobin polymorphism in breast cancer patients from Jordan. *Clinica chimica acta; international journal of clinical chemistry* 2004; **341**(1-2): 17-21.
9. Ubaidullaev AM, Kazakov KS, Khakimov MA. [Genetic variants of haptoglobin in patients with nephrotuberculosis]. *Problemy tuberkuleza* 2002; (7): 42-3.
10. Bjornsson E, Helgason H, Halldorsson G, et al. A rare splice donor mutation in the haptoglobin gene associates with blood lipid levels and coronary artery disease. *Human molecular genetics* 2017; **26**(12): 2364-76.

11. Bulters D, Gastra B, Zolnourian A, et al. Haemoglobin scavenging in intracranial bleeding: biology and clinical implications. *Nature reviews Neurology* 2018.
12. Froguel P, Ndiaye NC, Bonnefond A, et al. A genome-wide association study identifies rs2000999 as a strong genetic determinant of circulating haptoglobin levels. *PloS one* 2012; **7**(3): e32327.
13. Soejima M, Sagata N, Komatsu N, et al. Genetic factors associated with serum haptoglobin level in a Japanese population. *Clinica chimica acta; international journal of clinical chemistry* 2014; **433**: 54-7.
14. Boyd A, Golding J, Macleod J, et al. Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2013; **42**(1): 111-27.
15. Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International journal of epidemiology* 2013; **42**(1): 97-110.
16. Visvikis-Siest S, Siest G. The STANISLAS Cohort: a 10-year follow-up of supposed healthy families. Gene-environment interactions, reference values and evaluation of biomarkers in prevention of cardiovascular diseases. *Clinical chemistry and laboratory medicine* 2008; **46**(6): 733-47.
17. Shahabi P, Siest G, Herbeth B, Ndiaye NC, Visvikis-Siest S. Clinical necessity of partitioning of human plasma haptoglobin reference intervals by recently-discovered rs2000999. *Clinica chimica acta; international journal of clinical chemistry* 2012; **413**(19-20): 1618-24.
18. Langlois MR, Delanghe JR. Biological and clinical significance of haptoglobin polymorphism in humans. *Clinical chemistry* 1996; **42**(10): 1589-600.