bjh short report

HbA_2 levels in normal adults are influenced by two distinct genetic mechanisms

Stephan Menzel,[1]* Chad Garner,[2]* Helen Rooks,[1] Tim D. Spector[3] and Swee Lay Thein[1,4]

[1]King's College London, Molecular Haematology, James Black Centre, London, UK, [2]Department of Epidemiology, School of Medicine, University of California Irvine, Irvine, CA, USA, [3]Twin Research & Genetic Epidemiology, King's College London, and [4]King's College Hospital NHS Foundation Trust, Department of Haematological Medicine, London, UK

Received 11 July 2012; accepted for publication 27 August 2012

Correspondence: Swee Lay Thein, Molecular Haematology, King's College London, James Black Centre, 125 Coldharbour Lane, London SE5 9NU, UK.
E-mail: sl.thein@kcl.ac.uk

*Both authors contributed equally.

Summary

Using a genome-wide association study, we found that common inter-individual differences in haemoglobin A_2 (HbA_2, $\alpha_2\delta_2$) levels are largely governed by genetic factors (42% of variability). The influence of age (1%) and sex (4%) was small. HbA_2 levels were influenced by two loci: the *HBS1L-MYB* locus on chromosome 6q, which has been shown to have pleiotropic effects on other haematological traits; and a second locus surrounding *HBB*, the gene encoding β-globin. Our results suggest that HbA_2 levels in adults are influenced by two different biological processes: one via kinetics of erythropoiesis, and the other, via competition between *HBB* and *HBD* activity.

Keywords: haemoglobin regulation, thalassaemias, globin genes, genetics.

Haemoglobin A_2 (HbA_2, $\alpha_2\delta_2$) is a minor adult haemoglobin, comprising between 2·0% and 3·2% of the total circulating haemoglobin in healthy adults (Weatherall & Clegg, 2001). At the molecular level, a single gene (*HBD*) encodes δ-globin and is found in a cluster on chromosome 11 with the other genes encoding the β–like globins. The genes are arranged in the order they are expressed during development and are sequentially activated as the erythroid progenitors differentiate terminally in adult haematopoiesis. Studies have suggested that molecular switches between embryonic (*HBE1*, ε), fetal (*HBG1/2*, γ) and adult (*HBD/HBB*, δ/β) genes rely on competition between the globin promoters for access to the upstream regulatory elements (β locus control region) (Palstra *et al*, 2008). During this process, the gene closest to the β locus control region (β–LCR), i.e. *HBE1*, is activated first and *HBB*, the one furthest away, is activated last, but autonomous silencing of the preceding gene also contributes to the haemoglobin switching process. The ability to compete for the β–LCR and autonomous silencing rely on the change in the mediating transcription factors, some activating, and others, repressing.

It is presently not known how HbA_2 (and δ-globin chain) expression changes during adult erythropoiesis, but it has been suggested that δ-globin chain synthesis declines as maturation in erythroid progenitors progresses (Steinberg & Nagel, 2009). As *HBD* is located between the two γ (*HBG1*, *HBG2*) and β (*HBB*) genes, one can surmise that its expression during erythroid maturation closely follows that of γ, but precedes that of β, and peaks before progenitors reach full maturity. Earlier erythroid progenitors produce significant amounts of fetal haemoglobin (HbF, $\alpha_2\gamma_2$), while more mature progenitors contain nearly 100% adult haemoglobin and little, if any, HbF (Wojda *et al*, 2002). F cells (FC) represent a minor sub-population of erythrocytes that contain most of the fetal haemoglobin detectable in peripheral blood. They are thought to derive from less mature erythroid progenitors and are more numerous during situations of acute haemoglobin demand ('stress erythropoiesis'). We have previously reported that FC percentage in healthy adults is largely genetically determined, with three major acting loci: the β-globin locus, *HBSL1-MYB* and *BCL11A* (Menzel *et al*, 2007a).

 doi:10.1111/bjh.12084

HbA_2 has no known physiological function, but measurement of HbA_2 values is essential in screening programmes for thalassaemia (Weatherall & Clegg, 2001). Elevated HbA_2 levels in the presence of hypochromic microcytic red blood cells is diagnostic of heterozygous β-thalassaemia, and results in part from increased transcriptional activity from both *cis* and *trans HBD* genes (Codrington *et al*, 1990). Generally, HbA_2 values tend to be higher in association with severe β-thalassaemia mutations, with one exception. It has been observed that heterozygotes for mutations that affect the *HBB* promoter (very mild or null) have unusually high HbA_2 and HbF values (Huisman, 1997). In this case, inactivation of the *HBB* promoter removes competition for the upstream β–LCR and limiting transcription factors, allowing greater interaction with the *HBD* and *HBG* promoters leading to increased HbA_2 and HbF levels. Some of the increased HbA_2 levels in β-thalassaemia carriers are post-translational, reflecting the β-globin deficit and increased availability of α-globin. Low HbA_2 levels result from reduced synthesis of δ-globin chains (as in δ-thalassaemia) or from a post-translational mechanism due to insufficient α-globin (as in α-thalassaemia or iron deficiency).

These observations illustrate the complexity of HbA_2 control. Nonetheless, earlier studies of the intra-familial segregation of HbA_2 values in heterozygous β-thalassaemia suggest that genetic factors account for some of the common variation in HbA_2 levels in adults (Weatherall & Clegg, 2001). We have used a genome-wide association study (GWAS) to investigate the genetic factors contributing to HbA_2 variability in healthy European adults, and the relationship between HbA_2 and FC.

Subjects and methods

Healthy Northern Europeans from the Twins UK British twin registry had previously undergone genome-wide single nucleotide polymorphism (SNP) typing with Illumina Infinium technology (La Jolla, CA, USA), generating two groups of genotyped twins: a study group (n = 2322) (extracted from TwinsUK-II), and a replication group (n = 1716) (extracted from TwinsUK-I). In both groups, HbA_2 (% of total haemoglobin) was measured by high performance liquid chromatography and FC (% of erythrocytes containing fetal haemoglobin) by flow cytometry. Linear regression modelling and linear mixed effects modelling was used to study the effect of SNP genotypes and covariates on HbA_2 and FC levels. Heritability was estimated with the structural equation modelling program Mx (http://www.vcu.edu/mx/) using the observed HbA_2 covariance matrices for the monozygotic and dizygotic twin pairs (see also supplementary information).

Results and discussion

HbA_2% was found to be normally distributed (average 2·73 ± 0·27 standard deviation, n = 2340). Minimum and

Table I. Results for the genome-wide association study (GWAS) of HbA_2 levels.

			HbA_2 analysis						
			GWAS		Replication	F cells analysis		HbA_2 analysis accounting for F cells	
SNP	Chromosome position	Minor allele (frequency)	*P*-value	R^2	*P*-value	*P*-value	R^2	*P*-value	R^2
Chromosome 6q									
rs1547247	135432529	A (0·31)	$3{\cdot}02 \times 10^{-8}$	0·01	$1{\cdot}10 \times 10^{-5}$	$1{\cdot}41 \times 10^{-30}$	0·05	$1{\cdot}39 \times 10^{-5}$	0·01
rs7775698	135460328	T (0·26)	$2{\cdot}51 \times 10^{-9}$	0·02	n.a.	$1{\cdot}76 \times 10^{-96}$	0·15	$3{\cdot}30 \times 10^{-5}$	0·01
rs9399137	135460711	C (0·26)	$5{\cdot}19 \times 10^{-9}$	0·02	$3{\cdot}90 \times 10^{-9}$	$2{\cdot}81 \times 10^{-92}$	0·14	$4{\cdot}50 \times 10^{-5}$	0·01
rs4895441	135468266	G (0·27)	$5{\cdot}72 \times 10^{-9}$	0·02	n.a.	$1{\cdot}10 \times 10^{-92}$	0·14	$5{\cdot}50 \times 10^{-5}$	0·01
rs9376092	135468837	A (0·27)	$8{\cdot}69 \times 10^{-9}$	0·01	$2{\cdot}82 \times 10^{-10}$	$1{\cdot}48 \times 10^{-90}$	0·14	$7{\cdot}16 \times 10^{-5}$	0·01
rs9494145	135474245	C (0·23)	$1{\cdot}03 \times 10^{-8}$	0·01	$1{\cdot}62 \times 10^{-8}$	$2{\cdot}33 \times 10^{-70}$	0·11	$1{\cdot}42 \times 10^{-5}$	0·01
Chromosome 11p									
rs11036212	5178401	G (0·17)	$1{\cdot}45 \times 10^{-9}$	0·02	$1{\cdot}65 \times 10^{-10}$	0·0055	0·004	$1{\cdot}03 \times 10^{-8}$	0·01
rs7950726	5182023	A (0·12)	$1{\cdot}23 \times 10^{-11}$	0·02	$1{\cdot}17 \times 10^{-11}$	0·0028	0·004	$1{\cdot}99 \times 10^{-10}$	0·02
rs12787404	5182342	A (0·12)	$6{\cdot}82 \times 10^{-12}$	0·02	n.a.	0·0049	0·004	$1{\cdot}01 \times 10^{-10}$	0·02
rs10837582	5185284	G (0·35)	$3{\cdot}12 \times 10^{-8}$	0·01	n.a.	0·59	<0·001	$7{\cdot}23 \times 10^{-8}$	0·01
rs12793110	5188141	T (0·12)	$5{\cdot}11 \times 10^{-12}$	0·02	n.a.	0·0047	0·004	$7{\cdot}73 \times 10^{-11}$	0·02
rs10837628	5200980	G (0·34)	$1{\cdot}53 \times 10^{-9}$	0·02	n.a.	0·92	<0·001	$2{\cdot}41 \times 10^{-9}$	0·02
rs11036364	5205580	A (0·50)	$2{\cdot}80 \times 10^{-9}$	0·01	n.a.	0·034	0·001	$2{\cdot}59 \times 10^{-8}$	0·01

Shown are all single nucleotide polymorphisms (SNPs) that reached genome-wide significant association with HbA_2 in the GWAS study group (n = 2322). Genotypes for six of these SNPs were also available from the replication group (n = 1716) and the resulting *P*-values are shown. QTL11.1 is represented by an LD block of rs11036212, rs7950726, rs12787404 and rs12793110; QTL11.2 includes rs10837582 and rs10837628; QTL11.3 is rs11036364. R^2 represents the proportion of the overall population variance that can be attributed to the SNP.

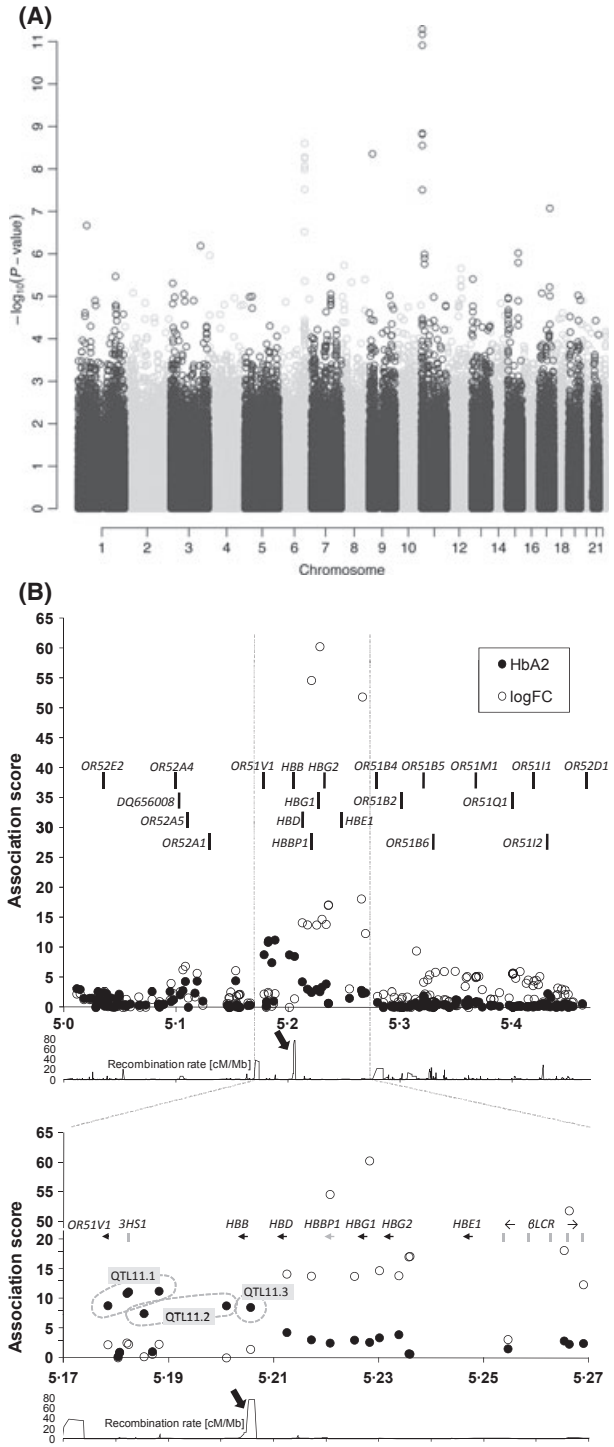

Fig 1. Association of SNPs with HbA_2 levels. (A) Genome-wide association scores (−log of *P*-value), showing two regions of association, one at chromosome 6q23.3 and the other, at 11p15.4. (B) HbA_2 association at chromosome 11p15.4 with the β-globin gene cluster at its centre. Association scores for F cell values (open circles) are shown for comparison. Physically and genetically, the gene cluster is split into two regions. On the left side (to about 5·206 Mbp), the downstream region of the-globin gene cluster contains the 3′HS regulatory site (Fleenor and Kaufman, 1993), the β-globin gene (*HBB*) itself and also the SNPs significantly associated with HbA_2 levels. On the right side of the graph, the larger part of the cluster contains the remaining β-like genes (*HBD*, *HBG1, HBG2* and *HBE1*) and the distant upstream regulatory region, the β–LCR. This region also contains a major association signal for F cells (FC) (Garner *et al*, 2000) and HbF (Uda *et al*, 2008). Both regions are separated by a well-characterized recombination hotspot (Holloway *et al*, 2006) (indicated by an arrow) just upstream of *HBB*, with its centre approximately at position chr 11: 5 205 880 (NCBI 36; http://genome.ucsc.edu). All significant HbA_2-associated SNPs reside to the left of the hotspot, with the exception of rs11036364, which is inside, but to the left of the recombination maximum. HbA_2-associated SNPs exist in three largely independent loci, QTL11.1, QTL11.2, QTL11.3 (dotted circles), all clustered around *HBB* and near *HBB* regulatory elements. The European recombination rate was plotted from HapMap data (http://hapmap.ncbi.nlm.nih.gov/downloads/index.html.en) and other features of the graph were localized using the UCSC genome browser (NCBI36) at http://genome.ucsc.edu.

maximum values were 1·30 and 3·70, respectively, after seven extreme outliers (>4 standard deviations outside the mean) had been removed. Genetic factors (heritability) comprise 42% (95% confidence interval = 12–63%) of the trait variance. Whilst females retained more HbF than males, the opposite was true for HbA_2 (P-value = $4{\cdot}07 \times 10^{-6}$, $R^2 = 0{\cdot}01$). Similar to HbF, HbA_2 declined with age (P-value = $3{\cdot}54 \times 10^{-20}$, $R^2 = 0{\cdot}04$). HbA_2 levels showed a weak, but statistically significant positive correlation of 0·14 with FC levels, P-value < 0·01). Therefore, FC must contain, on average, more HbA_2 than the rest of the erythrocytes. More speculatively, a population of 'A_2 cells' might exist, analogous to FC, and both of these red cell populations might overlap partially.

Genome-wide association analysis was performed with 531 038 SNPs in the study group of 2322 individuals. Thirteen SNPs in two independent genomic regions, showed genome-wide statistically significant associations with HbA_2 (Table I, Fig 1A): six SNPs spanning approximately 42 kb on chromosome 6q23.3 (at *HBS1L-MYB*, block 2) (Thein *et al*, 2007) and seven SNPs spanning about 27 kb on 11p15.4 (at the β-globin locus, Fig 1B). The combined loci account for 5·5% of the total variance in HbA_2 (2% and 3·5%, respectively) and approximately 13% (5·5/42%) of the heritability.

Stepwise regression indicated that a single genetic association was responsible for all six individual SNP associations in the *HBS1L-MYB* signal, with the strongest signal at rs7775698 (P-value = $2{\cdot}51 \times 10^{-9}$) accounting for the effects of the remaining five SNPs. This marker and three others (rs9399137, rs4895441 and rs9376092) form a block of strong linkage disequilibrium (LD) within Northern European populations (LD, pairwise $R^2 > 0{\cdot}90$), which was previously reported as a major determinant of FC levels (HMIP-2) (Thein *et al*, 2007) and a modulator of diverse haematological parameters (Menzel *et al*, 2007b; Soranzo *et al*, 2009).

Individuals with the minor (less common) allele at this locus have more HbA_2, but also substantially more FC (Thein *et al*, 2007), and marginally fewer, but larger red blood cells that carry more haemoglobin, more platelets, fewer monocytes and fewer white cells (Menzel *et al*, 2007b; Soranzo *et al*, 2009). Statistically adjusting for the effect of FC levels considerably reduced the strength of the HbA_2 association (Table I) but a statistically significant effect of the locus remained. Our interpretation is that the 6q23.3 locus affects HbA_2 levels, at least partially, through altering the kinetics of haematopoiesis and promoting FC (having more HbA_2) and possibly 'A_2 cells'.

Detailed examination of the associated SNPs at the β-globin locus suggests the presence of three independent association signals (Fig 1B), all situated around the *HBB*, which encodes β-globin. QTL11.1 is represented by an LD block of four SNPs (rs11036212, rs7950726, rs12787404 and rs12793110, strongest association P-value = $5{\cdot}11 \times 10^{-12}$), which covers a 10-kb region downstream of *HBB* and encompasses the regulatory 3′ DNase hypersensitive site (Forget & Hardison, 2009). QTL11·2 (rs10837582 and rs10837628), also downstream of *HBB*, is possibly linked to its downstream proximal enhancer (Forget & Hardison, 2009). QTL11.3 (rs11036364) is 703 nucleotides upstream of the *HBB* transcription start site, adjacent to its promoter (Forget & Hardison, 2009). The region containing QTLs 11.1, 11.2 and 11.3 is physically and genetically (Holloway *et al*, 2006) distinct from the region of *HBD*, the actual gene encoding the δ-globin chain, and also from that of major FC and HbF association (Garner *et al*, 2000; Menzel *et al*, 2007a; Uda *et al*, 2008) at chr11p15.4 (Fig 1B). Thus, it is genetic variation in the vicinity of the sequence encoding β-globin (*HBB*), not δ-globin (*HBD*) that ultimately determines relative HbA_2 abundance.

In contrast to the HMIP-2 locus on chromosome 6, the chromosome 11 locus contributes very little to FC variability (Table I), and this is largely restricted to QTL11.1. Also in contrast to HMIP-2 (Menzel *et al*, 2007b; Soranzo *et al*, 2009), no effect of QTL11 alleles on MCV and MCH can be detected, even in an extended dataset of 13 943 individuals (Table SI). Our results are substantially different from those reported for a Sardinian population (Uda *et al*, 2008), in which the main association with HbA_2 levels was detected at the site of the β–LCR with a weak association at the alpha-globin locus on chromosome 16, neither of which were present in our data. However, thalassaemic variants might be present in the Sardinian cohort.

We detected two loci that illustrate two principally different biological processes modulating peripheral HbA_2 levels in healthy adults. We believe that the *HBS1L-MYB* locus on chromosome 6 affects HbA_2 levels predominantly through the kinetics of erythropoiesis. Our results suggest that FC contain more HbA_2 than non–F erythrocytes, supporting the notion that the HbA_2 content and Hb composition of erythroid cells change during maturation (Steinberg & Nagel, 2009). It seems that the second locus on chromosome 11p affects *HBD* expression via competitive interaction with *HBB*, simulating a mild β-thalassaemia phenotype consistent with the proposed effect of rare *KLF1* variants on HbA_2 expression (Perseu *et al*, 2011). However, we found no evidence of any quantitative effect from common genetic *KLF1* variation (on chromosome 19p) on HbA_2 in this GWAS.

Acknowledgements

We thank Nicole Soranzo and John Chambers from Haem-Gen Consortium for performing the look-up on our HbA_2-associated SNPs on chromosome 11. We thank Claire Steward for help in preparation of the manuscript. This work was supported by the Medical Research Council, UK (Grant G0000111, ID51640 to SLT). Twins UK (TUK): The study was funded by the Wellcome Trust and ENGAGE project grant agreement HEALTH-F4-2007-201413. The study also receives support from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London. TDS is an NIHR senior Investigator and ERC Senior Researcher.

Authorship contribution

SM designed and conducted the study, interpreted the data and wrote the manuscript; CG performed statistical and genetic analysis, and wrote the manuscript; HR performed genotyping and interpreted data; TS contributed valuable material; SLT designed the study, interpreted data and wrote the manuscript.

Conflict of interest disclosure

All authors declare no competing interests.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig S1. Distribution of HbA_2 (% of total Hb) in an adult normal Northern European cohort (TwinsUK).

Table SI. No association of HbA_2-associated SNPs with erythrocyte hemoglobin content (MCH) or volume (MCV) in 13 943 healthy European individuals.

Data S1 Supplemental Methods.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

References

Codrington, J.F., Li, H.-W., Kutlar, F., Gu, L.-H., Ramachandran, M. & Huisman, T.H.J. (1990) Observations on the levels of Hb A_2 in patients with different b-thalassemia mutations and a d chain variant. *Blood*, **76**, 1246–1249.

Fleenor, E.F. & Kaufman, R.E. (1993) Characterization of the DNase I hypersensitive site 3′ of the human beta globin gene domain. *Blood*, **10**, 2781–2790.

Forget, B.G. & Hardison, R.C. (2009) The normal structure and regulation of human globin gene clusters. In: Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management, 2nd edn (ed. by M.H. Steinberg, B.G. Forget, D.R. Higgs & D.J. Weatherall), pp. 46–61. Cambridge University Press, New York, USA.

Garner, C., Tatu, T., Game, L., Cardon, L.R., Spector, T.D., Farrall, M. & Thein, S.L. (2000) A candidate gene study of F cell levels in sibling pairs using a joint linkage and association analysis. *GeneScreen*, **1**, 9–14.

Holloway, K., Lawson, V.E. & Jeffreys, A.J. (2006) Allelic recombination and de novo deletions in sperm in the human beta-globin gene region. *Human Molecular Genetics*, **15**, 1099–1111.

Huisman, T.H. (1997) Levels of Hb A2 in heterozygotes and homozygotes for beta-thalassemia mutations: influence of mutations in the CAC-CC and ATAAA motifs of the beta-globin gene promoter. *Acta Haematologica*, **98**, 187–194.

Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., Foglio, M., Zelenika, D., Boland, A., Rooks, H., Best, S., Spector, T.D., Farrall, M., Lathrop, M. & Thein, S.L. (2007a) A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15 . *Nature Genetics*, **39**, 1197–1199.

Menzel, S., Jiang, J., Silver, N., Gallagher, J., Cunningham, J., Surdulescu, G., Lathrop, M., Farrall, M., Spector, T.D. & Thein, S.L. (2007b) The HBS1L-MYB intergenic region on chromosome 6q23.3 influences erythrocyte, platelet, and monocyte counts in humans. *Blood*, **110**, 3624–3626.

Palstra, R.J., de Laat, W. & Grosveld, F. (2008) Beta-globin regulation and long-range interactions. *Advances in Genetics*, **61**, 107–142.

Perseu, L., Satta, S., Moi, P., Demartis, F.R., Manunza, L., Sollaino, M.C., Barella, S., Cao, A. & Galanello, R. (2011) KLF1 gene mutations cause borderline HbA2. *Blood*, **118**, 4454–4458.

Soranzo, N., Spector, T.D., Mangino, M., Kuhnel, B., Rendon, A., Teumer, A., Willenborg, C., Wright, B., Chen, L., Li, M., Salo, P., Voight, B. F., Burns, P., Laskowski, R.A., Xue, Y., Menzel, S., Altshuler, D., Bradley, J.R., Bumpstead, S., Burnett, M.S., Devaney, J., Doring, A., Elosua, R., Epstein, S.E., Erber, W., Falchi, M., Garner, S.F., Ghori, M.J., Goodall, A.H., Gwilliam, R., Hakonarson, H.H., Hall, A.S., Hammond, N., Hengstenberg, C., Illig, T., Konig, I.R., Knouff, C.W., McPherson, R., Melander, O., Mooser, V., Nauck, M., Nieminen, M.S., O'Donnell, C.J., Peltonen, L., Potter, S.C., Prokisch, H., Rader, D.J., Rice, C.M., Roberts, R., Salomaa, V., Sambrook, J., Schreiber, S., Schunkert, H., Schwartz, S.M., Serbanovic-Canic, J., Sinisalo, J., Siscovick, D.S., Stark, K., Surakka, I., Stephens, J., Thompson, J.R., Volker, U., Volzke, H., Watkins, N.A., Wells, G.A., Wichmann, H.E., Van Heel, D.A., Tyler-Smith, C., Thein, S.L., Kathiresan, S., Perola, M., Reilly, M.P., Stewart, A.F., Erdmann, J., Samani, N.J., Meisinger, C., Greinacher, A., Deloukas, P., Ouwehand, W.H. & Gieger, C. (2009) A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nature Genetics*, **41**, 1182–1190.

Steinberg, M.H. & Nagel, R.L. (2009) Hemoglobins of the embryo, fetus, and adult. In: Disorders of Hemoglobin: Genetics, Pathophysiology, and Clinical Management, 2nd edn (ed. by M.H. Steinberg, B.G. Forget, D.R. Higgs & D.J. Weatherall), pp. 119–135. Cambridge University Press, New York, USA.

Thein, S.L., Menzel, S., Peng, X., Best, S., Jiang, J., Close, J., Silver, N., Gerovasilli, A., Ping, C., Yamaguchi, M., Wahlberg, K., Ulug, P., Spector, T.D., Garner, C., Matsuda, F., Farrall, M. & Lathrop, M. (2007) Intergenic variants of HBS1L-MYB are responsible for a major quantitative trait locus on chromosome 6q23 influencing fetal hemoglobin levels in adults. *Proceedings of the National Academy of Sciences of the USA*, **104**, 11346–11351.

Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V.G., Chen, W., Usala, G., Busonero, F., Maschio, A., Albai, G., Piras, M.G., Sestu, N., Lai, S., Dei, M., Mulas, A., Crisponi, L., Naitza, S., Asunis, I., Deiana, M., Nagaraja, R., Perseu, L., Satta, S., Cipollina, M.D., Sollaino, C., Moi, P., Hirschhorn, J.N., Orkin, S.H., Abecasis, G.R., Schlessinger, D. & Cao, A. (2008) Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. *Proceedings of the National Academy of Sciences of the USA*, **105**, 1620–1625.

Weatherall, D.J. & Clegg, J.B. (2001) The Thalassaemia Syndromes. Blackwell Science, Oxford.

Wojda, U., Noel, P. & Miller, J.L. (2002) Fetal and adult hemoglobin production during adult erythropoiesis: coordinate expression correlates with cell proliferation. *Blood*, **99**, 3005–3013.