Human Molecular Genetics, 2009, Vol. 18, No. 14 **2711–2718**
doi:10.1093/hmg/ddp203
Advance Access published on May 6, 2009

Common variants in the *SLCO1B3* locus are associated with bilirubin levels and unconjugated hyperbilirubinemia

Serena Sanna[1,†], Fabio Busonero[1,†], Andrea Maschio[1], Patrick F. McArdle[2], Gianluca Usala[1], Mariano Dei[1], Sandra Lai[1], Antonella Mulas[1], Maria Grazia Piras[1], Lucia Perseu[1], Marco Masala[1], Mara Marongiu[1], Laura Crisponi[1], Silvia Naitza[1], Renzo Galanello[3], Gonçalo R. Abecasis[4], Alan R. Shuldiner[2,5], David Schlessinger[6], Antonio Cao[1] and Manuela Uda[1,*]

[1]Istituto di Neurogenetica e Neurofarmacologia del Consiglio Nazionale delle Ricerche, Monserrato, 09042 Cagliari, Italy, [2]Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, MD, USA, [3]Clinica Pediatrica, Ospedale Regionale delle Microcitemie, Dipartimento di Scienze Biomediche e Biotecnologie, Università degli Studi di Cagliari, 09121 Cagliari, Italy, [4]Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109, USA, [5]Veterans Administration Medical Center, Geriatric Research and Education Clinical Center, Baltimore, MD, USA and [6]Laboratory of Genetics, National Institute on Aging, Baltimore, MD, USA

Received February 5, 2009; Revised March 24, 2009; Accepted April 28, 2009

Bilirubin, resulting largely from the turnover of hemoglobin, is found in the plasma in two main forms: unconjugated or conjugated with glucuronic acid. Unconjugated bilirubin is transported into hepatocytes. There, it is glucuronidated by *UGT1A1* and secreted into the bile canaliculi. We report a genome wide association scan in 4300 Sardinian individuals for total serum bilirubin levels. In addition to the two known loci previously involved in the regulation of bilirubin levels, *UGT1A1* ($P = 6.2 \times 10^{-62}$) and *G6PD* ($P = 2.5 \times 10^{-8}$), we observed a strong association on chromosome 12 within the *SLCO1B3* gene ($P = 3.9 \times 10^{-9}$). Our findings were replicated in an independent sample of 1860 Sardinians and in 832 subjects from the Old Order Amish (combined $P < 5 \times 10^{-14}$). We also show that *SLC01B3* variants contribute to idiopathic mild unconjugated hyperbilirubinemia. Thus, *SLC01B3* appears to be involved in the regulation of serum bilirubin levels in healthy individuals and in some bilirubin-related disorders that are only partially explained by other known gene variants.

INTRODUCTION

Bilirubin results from hemoglobin catabolism (and to a lesser extent from other heme-containing proteins) within the reticulo-endothelial system. Bilirubin is transported into the liver, where specific carrier proteins belonging to the Organic Anion Protein Transporter family (1) are thought to mediate its uptake. In hepatocytes, bilirubin is glucuronidated by (UDP)-glucuronosyltransferase and thereby transformed into a water soluble conjugated molecule (2). Conjugated bilirubin is actively secreted into the bile canaliculi by a membrane ATP-dependent transporter, designated multidrug resistance-associated protein 2 (MRP2) (3). In plasma, bilirubin occurs primarily unconjugated tightly bound to albumin, or conjugated as a water-soluble glucuronide. High levels of unconjugated bilirubin in plasma may indicate increased production by hemolysis, ineffective erythropoiesis or decreased conjugation, whereas conjugated hyperbilirubinemia reflects decreased excretion by damaged biliary duct or hepatic parenchymal cells.

*To whom correspondence should be address. Tel: +39 0706754591; Fax: +39 0706754652; Email: manuela.uda@inn.cnr.it
†S.S. and F.B. equally contributed to this work.

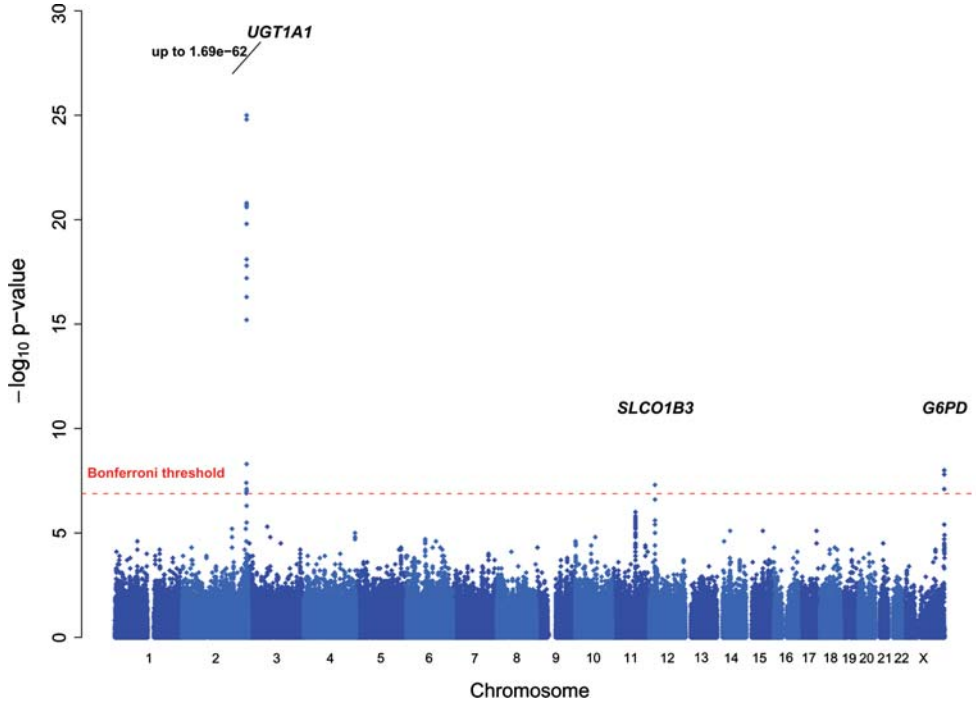

Figure 1. Genome-wide scan of serum total bilirubin. The figure summarizes association *P*-values (additive test) for all SNPs that passed quality control tests ($N = 362{,}129$). Position of *UGT1A*, *SLCO1B3* and *G6PD* genes are annotated. Red dotted line marks Bonferroni threshold (1.3×10^{-7}).

RESULTS

Within the SardiNIA study we conducted a genome-wide association (GWA) analysis (4,5) in 4300 Sardinians to identify genetic factors affecting a series of quantitative traits, including serum bilirubin levels (see Material and Methods). The strongest associations were observed on chromosome 2, in the promoter of *UGT1A1* (uridine diphosphoglucuronyltransferase) gene (rs887829, $P = 6.21 \times 10^{-62}$, Fig. 1 and Supplementary Material, Table S1), a major and well-known modifier of bilirubin levels (6), and on the X chromosome near the *G6PD* (glucose-6-phosphatedehydrogenase) gene (rs766420, $P = 9.40 \times 10^{-09}$, Fig. 1 and Supplementary Material, Table S1).

In addition, a third locus on chromosome 12p12.2 reached genome-wide significance (Fig. 1). The strongest *P*-value was observed at SNP rs2117032 ($P = 4.74 \times 10^{-08}$, Supplementary Material, Table S1) downstream of the 3′-UTR of the *SLCO1B3* (solute carrier organic anion transporter family, member 1B3) gene (7) (Fig. 2); each copy of the minor allele C was responsible, on average, for an increase of 0.048 mg/dl in bilirubin levels. The second ranked marker in this region was rs17680137 ($P = 2.3 \times 10^{-7}$), located in intron 7 of *SLCO1B3*, in modest LD with rs2117032 ($r^2 = 0.29$), and responsible for a slightly larger effect (0.067 mg/dl per copy of the minor allele G). Our genotype strategy was designed to take advantage of the relatedness of the sample. In fact we genotyped first a small number of markers (10K GeneChip) in largest families to identify shared stretches of chromosome between individuals, and then genotyped with a denser map (the 500K chip) a subset of individuals, selected to be dispersed along the known pedigrees with an overlap of 436 individuals genotyped with both platforms. Next a within-family imputation method was used to predict missing genotypes (see Material and Methods). The two markers, rs17680137 and rs2117032, belong to the 500 K SNP array and 10 K SNP array, respectively, and thus were imputed the first in the majority of the samples, the second in only one-fourth. To evaluate the effect of the different number of genotyped and imputed samples, we directly genotyped both markers in individuals for whom only imputed genotypes were available. Using all available genotypes, the association test showed SNP rs17680137 as more representative of total bilirubin variation ($P = 3.9 \times 10^{-9}$, see Table 1 and Supplementary Material, Table S1). Interestingly, when including this marker in the model, SNP rs2117032 still showed some evidence for association ($P = 0.00050$). Furthermore, we observed a significant interaction between the two markers ($P = 1.28 \times 10^{-04}$). Using all available genotypes, we were able to infer the complete haplotypes of 3431 individuals for which the wet-genotype for both markers were available and we observed that only haplotypes carrying both positive or both negative alleles were significant (H_{GC} $P = 1.8 \times 10^{-07}$ and H_{CT} $P = 2.1 \times 10^{-07}$, see Table 2). Notably, we observed that the haplotype frequencies were similar to the minor allele frequency of the two variants (H_{GC} freq = 0.26 and H_{CT} freq = 0.43), and that each haplotype was highly correlated with one variant (H_{GC} and rs17680137 cor = 0.97, H_{CT} and rs2117032 cor = 0.98) (Table 2). Those results were also replicated in the Amish population, where the only two haplotypes associated with bilirubin levels were those carrying both positives or both negatives alleles. Each haplotype was highly correlated with one of the two markers (Table 2). Thus the two SNPs, rather than having independent effects, are likely imperfectly tagging a third common variant of which the allele responsible for an increase in bilirubin levels would be probably associated with the H_{GC} haplotype, and the other allele with the H_{CT} haplotype.

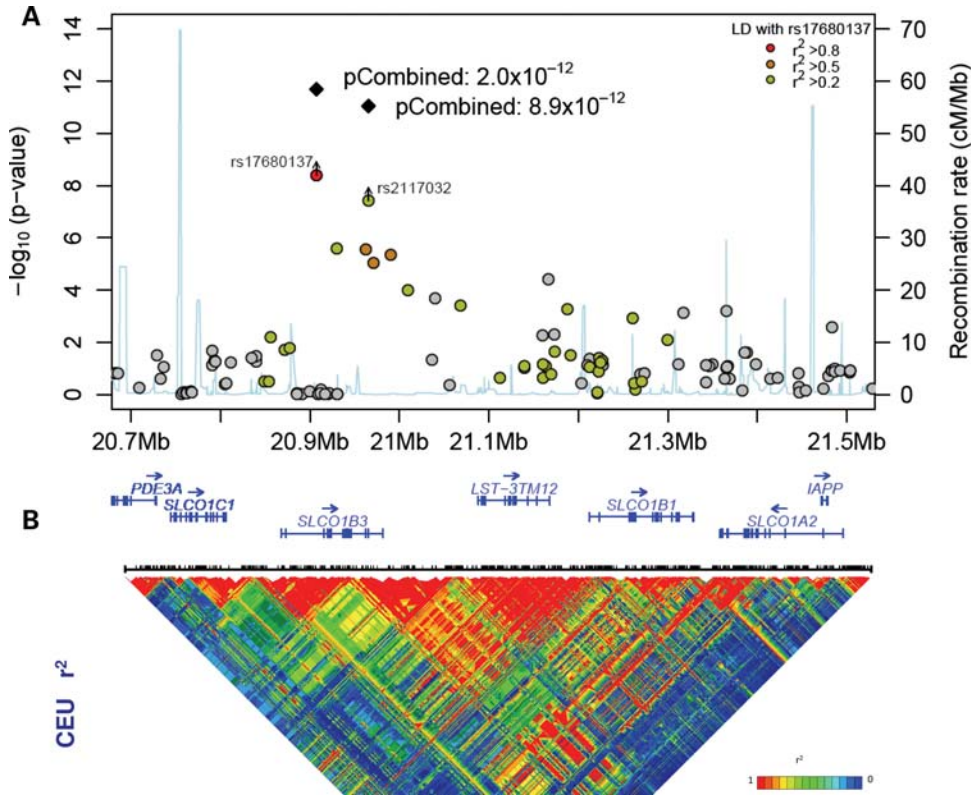

Figure 2. Evidence of (**A**) association with serum total bilirubin and (**B**) linkage disequilibrium around the *SLCO1B3* locus. (A) All SNPs in the *SLCO1B3* region (700 Kb) are plotted with association *P*-values (additive test) compared with physical position. Genomic position and gene annotations refer NCBI Build 36 map. SNPs are colored according to their linkage disequilibrium with the top variant (rs17680137). Blue line represents the recombination rate (in cM/Mb). Black diamonds represent combined *P*-values from SardiNIA and SardiNIA stage 2 cohorts. (B) Patterns of linkage disequilibrium (r^2) for the CEU HapMap population (Utah residents with European ancestry (16)) are plotted and colored as indicated in the palette.

We also evaluated the effect of possible confounding factors related to the particular hematological features of the Sardinian population, and repeated the association tests including as covariates Mean Corpuscular Hemoglobin, Mean Corpuscular Volume, total Hemoglobin, β-thalassemia carrier status and G6PD levels (8,9). None of those factors affected the association at the three loci. As expected, adjustment for G6PD levels reduced the association at the *G6PD* locus, somewhat strengthening the *UGT1A1* signal (Supplementary Material, Table S2). In addition, we observed that the association of *UGT1A1* and *SLCO1B3* appeared stronger in women than in men, although we found no significant difference in the magnitude of the effect between sexes. In contrast, the association with the X-linked *G6PD* locus was stronger in men, most likely because of the major impact of G6PD deficiency in males. In fact, there was no difference in association between *G6PD* genotype and bilirubin when G6PD level was taken into account (Supplementary Material, Table S3). Finally, none of the tests for epistasis between the top SNPs at the three loci was significant after correction for multiple testing (Supplementary Material, Table S4). Additional studies are required to confirm this observation.

Our findings at the *SLCO1B3* locus were successfully replicated in an independent sample of 1860 Sardinians, previously described and reported as SardiNIA stage 2 (10), with an effect in the same direction as that observed in the GWA sample ($P < 2 \times 10^{-04}$, Table 1). Moreover, we also replicated the association in 832 individuals enrolled in the Amish Heredity and Phenotype Intervention (HAPI) Heart study (11) ($P < 7 \times 10^{-03}$, Table 1), suggesting that variants in the *SLCO1B3* influence bilirubin levels not only in Sardinians but also in other populations of Caucasian origin. Combined, the evidence for association at this locus was $P < 5 \times 10^{-14}$ (Table 1).

As shown in Fig. 2, although our results point to the *SLCO1B3* gene, weaker association encompasses a genomic region of 700 Kb harboring, in addition to *SLCO1B3*, four more genes in the same family of membrane transport proteins. Among them, *SLCO1B1* (12) has been previously implicated in bilirubin liver uptake (13) and associated with bilirubin levels in Asian populations. In this gene two coding variants, namely Val174Ala/rs4149056 and Asp130Asn/rs2306283, have been described as major determinants of bilirubin variation (14,15), and they weakly correlated with rs17680137 ($r^2 = 0.008$ and $r^2 = 0.359$, respectively) and with rs2117032 ($r^2 = 0.031$ and $r^2 = 0.302$, respectively). To compare their effects with polymorphisms at *SLCO1B3* we used inferred genotypes based on HapMap (16) CEU (Utah residents with ancestry from northern and western Europe) haplotypes available from previous studies (17), and tested for association with total, conjugated and unconjugated bilirubin. Only nominal association with total bilirubin was detected at these coding variants ($P = 0.079$ and $P = 3.6 \times 10^{-03}$, Table 3). In addition, when including rs4149056 and

Table 1. Summary of association results for SNPs in the*SLCO1B3* locus in all cohorts

Marker	Study	N	Allele High/Low	Freq	Effect (StdErr)	Effect (StdErr) mg/dl	*P*-value
rs17680137	Stage 1						
	SardiNIA	4300	G/C	0.293	0.170 (0.029)	0.067 (0.011)	3.9×10^{-09}
	Stage 2						
	SardiNIA stage2	1860	G/C	0.273	0.143 (0.037)	0.044 (0.014)	1.2×10^{-04}
	Old Older Amish	832	G/C	0.229	0.171 (0.062)	0.040 (0.020)	6.1×10^{-03}
	Overall	6992					4.41×10^{-14}
rs2117032	Stage 1						
	SardiNIA	4300	C/T	0.466	0.129 (0.023)	0.048 (0.009)	3.8×10^{-08}
	Stage 2						
	SardiNIA stage2	1860	C/T	0.485	0.134 (0.033)	0.040 (0.012)	5.0×10^{-05}
	Old Older Amish[a]	838	C/T	0.415	0.170 (0.051)	0.034 (0.016)	5.2×10^{-04}
	Overall	6998					2.91×10^{-14}

The tables summarize the association results in all cohorts. For each marker, frequency and effect estimates are given with respect to the 'high frequency' allele. Association parameters (Effect, StdErr, *P*-value) are relative to the model where the trait is transformed, thus the effect estimated in this model represents the change in standard deviation units. Effect in mg/dl is also given for each marker. SardiNIA results are relative to the imputed genotypes inferred from Affymetrix and TaqMan genotypes.
[a]Old Older Amish sample refers to SNP rs10770763, $r^2 = 0.96$ with rs2117032 in CEU HapMap population, for which additional 6 genotyped individuals were available.

Table 2. Association analysis of haplotypes resulting from all possible combinations of alleles at markers rs17680137 and rs2117032

Haplotype	rs17680137	rs2117032[a]	Freq	Effect[b]	StdErr	*P*-value	Correlation[c] with rs17680137	rs2117032[a]
SardiNIA								
H_{CC}	C	C	0.294	0.022	0.031	0.49	0.38	0.56
H_{CT}	C	T	0.434	−0.143	0.027	2.10×10^{-07}	0.59	0.98
H_{GC}	G	C	0.262	0.164	0.031	1.84×10^{-07}	0.97	0.58
H_{GT}	G	T	0.010	0.018	0.133	0.893	0.17	0.09
Old Older Amish								
H_{CC}	C	C	0.202	0.058	0.067	0.398	0.23	0.60
H_{CT}	C	T	0.590	−0.150	0.055	0.007	0.64	1.00
H_{GC}	G	C	0.208	0.162	0.068	0.016	1.00	0.64
H_{GT}	G	T	0.000	–	–	–	–	–

[a]For the Amish sample, rs10770763 (T/C alleles) was genotyped instead of rs2117032 (C/T alleles).
[b]Effect size and Standard Error are given in standard deviation units.
[c]Correlation is given in absolute value.

rs2306283 as covariates in the model, SNPs at the *SLCO1B3* locus still remain highly significant (Supplementary Material, Table S5), supporting the hypothesis that the association signal observed in the *SLCO1B3* is not modulated by an indirect effect of these coding variants. Furthermore, since the SNPs tested in the GWAS covered 66% of the common variants in the *SLCO1B1* at $r^2 > 0.5$, it is unlikely that other SNPs in this gene may have a stronger effect than those in the *SLCO1B3* locus. The differential contribution of the two transporters may be owing to population-specific genetic features. In Asians, for example, SNP rs17680137 is monomorphic.

Interestingly, we observed that SNPs in *SLCO1B1* were mostly associated with conjugated bilirubin ($P = 1.7 \times 10^{-04}$ for rs4149056), whereas no association was observed with unconjugated bilirubin ($P = 0.26$, Table 3). In contrast, SNP rs17680137 in *SLCO1B3* showed a stronger association with unconjugated bilirubin ($P = 5.2 \times 10^{-09}$) compared with conjugated bilirubin ($P = 0.0004$). A similar pattern of association was observed in the SardiNIA stage 2 sample (Table 3), where SNP rs4149056 was directly genotyped. Again, the association with unconjugated bilirubin at SNPs rs17680137 and rs2117032 remained significant after accounting for the two SNPs in the *SLCO1B1* gene (Supplementary Material, Table S5).

Conjugated bilirubin was not available in the Amish cohort, and thus a parallel analysis could not be performed.

We next tested for any effect of the *UGT1A1* gene on the association of bilirubin levels with the *SLCO1B3* gene. Including SNP rs887829, the top marker on the *UGT1A1* locus in the model, the association at rs2117032 and rs17680137 did not disappear but rather increased (data not shown), suggesting that *SLCO1B3* explains part of the variability not accounted for by *UGT1A1* polymorphisms. To explore this possibility further, we evaluated whether this locus may be responsible for idiopathic mild hyperbilirubinemia in patients lacking mutations in the *UGT1A1* gene. Hence, we genotyped 1004

Table 3. Association with total, conjugated and unconjugated bilirubin of SNPs in the *SLCO1B3* and *SLCO1B1* genes

Locus	Marker	Study	Freq	Total bilirubin Effect (StdErr)	*P*-value	Conjugated bilirubin Effect (StdErr)	*P*-value	Unconjugated bilirubin Effect (StdErr)	*P*-value
SLCO1B1	rs4149056 (C)	SardiNIA[a]	0.167	0.070 (0.040)	0.079	0.147 (0.039)	1.7×10^{-04}	0.045 (0.040)	0.26
		SardiNIA stage2	0.184	0.063 (0.043)	0.14	0.104 (0.042)	0.014	0.056 (0.043)	0.19
		Old Order Amish[b]	0.103	0.255 (0.086)	3.08×10^{-03}	–	Na	–	Na
			pCombined		0.0015		7.2×10^{-06}		0.095
SLCO1B1	rs2306283 (G)	SardiNIA[a]	0.485	0.090 (0.030)	3.6×10^{-03}	0.100 (0.03)	7.6×10^{-04}	0.086 (0.031)	0.0056
		SardiNIA stage2	–	–	Na	–	Na	–	Na
		Old Order Amish[c]	0.319	0.031 (0.054)	0.563	–	Na	–	Na
			pCombined		0.0039		7.6×10^{-04}		0.0056
SLCO1B3	rs17680137 (G)	SardiNIA	0.293	0.170 (0.029)	3.9×10^{-09}	0.100 (0.028)	0.0004	0.170 (0.029)	5.2×10^{-09}
		SardiNIA stage2	0.273	0.144 (0.037)	1.2×10^{-04}	0.107 (0.037)	0.0037	0.144 (0.037)	1.2×10^{-04}
		Old Order Amish	0.219	0.171 (0.062)	6.1×10^{-03}	–	Na	–	Na
			pCombined		4.41×10^{-14}		5.3×10^{-06}		2.7×10^{-12}

The table summarizes association results with total, conjugated and unconjugated bilirubin for SNP rs4149056 and SNP rs2306283, polymorphisms in the *SLCO1B1* gene previously known to be associated with neonatal hyperbilirubinemia. They are coding mutations, with the following aminoacid changes: Val174Ala and Asp130Asn, respectively. Effect size and Standard Error are given in standard deviation units. Conjugated and Unconjugated bilirubin were not available in the Amish cohort.

[a]SNPs rs4149056 and rs2306283 were imputed in all SardiNIA samples based on HapMap CEU haplotypes (see Material and Methods). Quality of imputation was $r^2 = 0.938$ and $r^2 = 0.929$, respectively.

[b]SNP rs4149056 was not genotyped in the Amish cohort, and rs11045879 ($r^2 = 0.86$) was analyzed as best proxy among genotyped SNPs.

[c]SNP rs2306283 was not genotyped in the Amish cohort, and rs2291075 ($r^2 = 0.86$) was analyzed as best proxy among genotyped SNPs.

Table 4. Association of *SLCO1B3* with hyperbilirubinemia

	N	C/C	C/G	G/G	C	G	Genotypic test *P*-value	Allelic test *P*-value
Healthy	1004	0.526	0.405	0.069	0.729	0.271		
Hyper-bilirubinemia	261	0.446	0.466	0.088	0.679	0.321	0.06	**0.029**
Hyper (TA6/TA6)	95	0.379	0.547	0.074	0.652	0.348	**0.019**	**0.031**
Hyper (TA6/TA7, TA7/TA7)	166	0.488	0.416	0.096	0.695	0.305	0.095	0.053

The table describes the allele and genotype frequencies in a data set of unrelated healthy (total bilirubin <1 mg/dl) and hyperbilirubinemia patients, screened for the *UGT1A1* A(TA)7TAA mutation. Hyperbilirubinemia was defined as total bilirubin >1.2 mg/dl. *P*-value <0.05 are highlighted in bold.

healthy individuals and 261 unrelated patients with hyperbilirubinemia (defined as serum total bilirubin >1.2 mg/dl) (2) from the SardiNIA sample for the mutation in the TATAA element of the 5′ promoter region of the *UGT1A1* gene (2). We observed a 5% enrichment in the frequency of the G allele of SNP rs17680137 in the hyperbilirubinemic patients compared with the healthy subgroup ($P = 0.029$; Table 4). Notably, most of the difference was attributable to patients homozygous for the wild-type allele, A(TA)6TAA at the *UGT1A* locus ($N = 95$, $P = 0.031$), rather than to patients carrying at least one copy of the mutated allele A(TA)7TAA ($N = 166$, $P = 0.053$) (2). Thus one can infer that *SLCO1B3* not only acts in the general population, but also affects mild hyperbilirubinemias.

DISCUSSION

GWA analysis for serum bilirubin levels on 4300 Sardinian revealed two striking signals in the *UGT1A1* and the *G6PD* genes. The association with *UGT1A1* is well known. In fact, mutations in this gene are responsible for the Crigler–Najjar syndrome as well as the more common mild unconjugated hyperbilirubinemia of Gilbert syndrome (2). The association of *G6PD* with bilirubin levels was also expected, given the high frequency of deficiency in G6PD in the Sardinian population; this condition is indeed associated with a shorter life span of red blood cells, resulting in a modest increase in bilirubin levels (18). Furthermore, the results revealed an additional locus on chromosome 2p12.2. The strongest associated marker in this region, an intronic SNP of the *SLCO1B3* gene, was successfully replicated in an independent Sardinian sample and in a cohort of Amish individuals. Notably, this gene is highly expressed in the basolateral membrane of hepatocytes and is a member of the OATP family encoding the Organic Anion Transporter Polypeptide Na^+-independent, OATP1B3 (7). The OATP proteins are expressed in a variety of tissues, including gut, brain, kidney and liver where they play important roles in drug absorption and distribution and excretion, with a wide spectrum of substrates that include both endogenous and xenobiotic compounds (1). Thus *SLCO1B3* is a plausible candidate gene responsible for

changes in bilirubin levels. Interestingly, another solute carrier organic anion transporter, *SLCO1B1*, previously implicated in liver bilirubin uptake (13), maps in the same region, but only weaker association was observed at this locus. In particular, we found that SNPs at *SLCO1B1* were mostly associated with conjugated bilirubin, although no association was detected with unconjugated bilirubin. Although we cannot exclude the possibility of a differential contribution of the two transporters in a population-specific fashion, the results produced in this study indicate that although the *SLCO1B3* gene regulates unconjugated serum bilirubin, *SLCO1B1* appears to act on conjugated bilirubin. *SLCO1B1*, with an high affinity for glucuronated bilirubin (19), may contribute (together with MPR3) to reverse transport from hepatocytes to plasma in normal conditions, and very likely also in cholestasis when the major export pump MPR2 is down-regulated (20).

Further studies are necessary to identify the *SLCO1B3*-specific causative variants and then to clarify the functional role of both OATP genes on conjugated and unconjugated bilirubin. Finally, the *SLCO1B3* SNPs identified here were also found to be associated with unconjugated hyperbilirubinemia of unknown origin. Speculatively, they may also act as modifiers of the clinical course of some phenotypes, as already shown for *UGT1A1* polymorphisms in hemolytic anemia, and anemia resulting from ineffective erythropoiesis, and neonatal jaundice (8,9,21). Thus, our findings not only reveal an additional regulator of bilirubin levels in the general population but may also contribute to the identification of genetic variants necessary to predict and prevent severe clinical manifestations in hematological and hepatic disorders linked to hyperbilirubinemia.

MATERIAL AND METHODS

Sample description

We recruited and phenotyped 6148 individuals, males and females, ages 14–102 years, from a cluster of four towns in the Ogliastra province of Sardinia (4). During physical examination, a blood sample was collected from each individual, and divided into two aliquots. One aliquot was used for genomic DNA extraction and the second (7 ml peripheral blood without anticoagulant) to characterize several blood phenotypes, including evaluation of serum total and conjugated bilirubin levels, measured with Express Plus (Bayer) equipment, according to manufacturer's instruction.

The normality range was 0.2–1.0 and 0–0.2 mg/dl for total and conjugated bilirubin, respectively. Hyperbilirubinemia was defined when total bilirubin serum concentration exceeded 1.2 mg/dl (2).

GWAS genotyping

We genotyped from the whole sample 4305 individuals selected to represent the largest available families, regardless of their phenotypic values (5). Specifically, 1412 were genotyped with the 500K Affymetrix Mapping Array Set and 3,329 with the 10K Mapping Array Set, with 436 individuals genotyped with both arrays. This genotyping strategy allowed us to examine the majority of our cohort in a cost-effective manner because genotypes for the SNPs that passed quality-control checks could be propagated through the pedigree via imputation (22,23). Bilirubin measurements were available for 4300 individuals among the 4305 genotyped. A total of 362 129 SNPs passed initial quality-control checks (5) and were tested for association. Although the statistical analysis was ongoing, an additional 2 million autosomal markers from HapMap imputed data became available. Repeating the GWA analysis using this larger data set of markers detected no additional loci, and the top markers at chromosomes 2 and 12 loci were confirmed. Further details of the strategy for imputation and data analysis are given in the Statistical Analysis section below.

Replication

We designed a ParAllele Custom Chip (from Affymetrix) to replicate the regions associated with bilirubin levels as well as other traits studied in the SardiNIA Project. Genotyping was performed in 1862 Sardinian, selected within the SardiNIA sample and independent from the individuals included in the GWA scan (kinship coefficient = 0 with those directly typed with the 500K chip), and indicated here as SardiNIA stage 2 data set (17). SNPs rs17680137 and rs2117032, were included in the ParAllele Custom chip, and also genotyped in additional individuals using TaqMan SNP genotyping assay (Applied Biosystems) according to the protocol. The concordance rate between inferred and wet genotypes was 74.7 and 93.4% for rs17680137 and rs2117032 (over $N = 1460$ and $N = 202$ samples), respectively. SNP rs4149056, not included in the custom chip was genotyped in the SardiNIA stage 2 data set using TaqMan technology. Mutations in the TATAA element of the 5′ promoter region of the *UGT1A1* gene were analyzed by direct sequencing (ABI-PRISM 377; Applied Biosystems, USA) of genomic DNA.

Statistical analysis

For individuals genotyped with a sparse map, we used a modified version of the Lander–Green algorithm (23) to estimate IBD sharing at the location of the SNPs being tested and identify stretches of haplotype shared with close relatives who were genotyped at higher density and probabilistically infer missing genotypes (22). Our inference approach allowed us to account for uncertainty in genotype assignment, estimating, instead of the most likely genotype, an expected genotype score, representing the expected number of copies of a reference allele (a fractional number between 0 and 2). The genotype scores were then used in the family-based association test for evaluating the additive effect of each marker, as described elsewhere (22,24). Owing to computational constraints, we divided large pedigrees into sub-units with 'bit-complexity' of 19 or less (typically, 20–25 individuals) before all the analyses. The association test fits a simple regression model and uses a variance component approach to account for correlation between different observed phenotypes within each family. To evaluate association on the X chromosome, we modeled a polygenic variance component shared according to an

X-linked kinship coefficient in addition to the usual autosomal polygenic variance component (4,24). Prior to all analyses, traits were normalized using quantile normalization to avoid inflation of type I error rates; gender, age and age squared were included as covariates.

In the SardiNIA stage 2 the association with bilirubin levels was tested as in the SardiNIA GWA sample. To infer haplotypes we used all available genotypes and Merlin software. We created a variable for each haplotype coded as 0, 1 or 2, representing the number of copies carried by an individual. A linear regression model was fitted for each haplotype. Epistasis and haplotype-specific tests were performed in the GWA sample using the R kinship package.

Old order Amish

The HAPI Heart Study was initiated in 2002. Participants of the HAPI Heart Study comprised adults from the Old Order Amish community of Lancaster County, PA, who were recruited over a 3-year period. The study aims and recruitment details, including ascertainment criteria, have been described previously (11). Physical examinations were conducted at the Amish Research Clinic in Strasburg, PA. A blood sample was obtained for measurement of bilirubin levels. Biochemical assays were performed by Quest Diagnostics (Horsham, PA, USA).

Genotypes for SNPs rs17680137, rs10770763 (proxy of rs2117032; $r^2 = 0.96$), rs11045879 (proxy of rs4149056; $r2 = 0.86$) and rs2291075 (proxy of rs2306283; $r^2 = 0.86$), were obtained from a previous GWAS performed using the 500K Affymetrix Mapping Array Set.

Statistical analysis was performed using in house developed software. In brief, we performed a measured genotype approach utilizing a *t*-test of the beta coefficient for the SNP variable. We included sex, age and age squared as fixed covariates in the model and a polygenic component modeled as a random effect to account for the full 14 generation pedigree of the Amish. The bilirubin trait was normalized using quantile normalization to avoid inflation of type I error rates; gender, age and age squared were included as covariates, as in the SardiNIA GWA sample.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We thank the many individuals who generously participated in this study, Monsignore Piseddu, Bishop of Ogliastra, the Mayors and citizens of the Sardinian towns (Lanusei, Ilbono, Arzana and Elini), and the head of the Public Health Unit ASL4, the province of Ogliastra for their volunteerism and cooperation. In addition, we are grateful to the Mayor and the administration in Lanusei for providing and furnishing the clinic site. We thank the team of physicians Angelo Scuteri, Marco Orrù, Maria Grazia Pilia, Maria Cristina Spada, Danilo Fois, Liana Ferreli, Francesco Loi, and nurses Paola Loi, Monica Lai and Anna Cau who carried out participant physical exams, and Monica Balloi who recruited the volunteers.

Conflict of Interest statement. None declared.

FUNDING

The work was supported by the following National Institutes of Health grants: HL084729 to G.R.A., HG002651 to G.R.A., National Institute of Aging contracts NO1-AG-1-2109 to the SardiNIA ('ProgeNIA') team and 263-MA-410953 to the University of Michigan (G.R.A).The Old Order Amish study was supported by National Institutes of Health research grant U01 HL72515, the University of Maryland General Clinical Research Center, grant M01 RR 16500, the Clinical Nutrition Research Unit of Maryland P30 DK072488, and the Baltimore Veterans Administration Medical Center Geriatrics Research.

AUTHOR CONTRIBUTIONS

G.R.A, D.S., A.C., A.R.S. and M.U. designed the study; R.G., A.S., D.S. and M.U. provided material and reagents; F.B., A.M., G.U., M.D., S.L. and A.M. performed the genotyping; M.G.P., L.P., M.M., M.M., A.R S. performed characterization of phenotype and data collection; S.S., P.F.M., analyzed the data; S.N., L.C., F.B., A.R.S., P.F.M. contributed to draft the manuscript; S.S., D.S., A.C. and M.U. wrote the paper.

REFERENCES

1. Hagenbuch, B. and Gui, C. (2008) Xenobiotic transporters of the human organic anion transporting polypeptides (OATP) family. *Xenobiotica*, **38**, 778–801.
2. Bosma, P.J., Chowdhury, J.R., Bakker, C., Gantla, S., de Boer, A., Oostra, B.A., Lindhout, D., Tytgat, G.N., Jansen, P.L., Oude Elferink, R.P. *et al.* (1995) The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in Gilbert's syndrome. *N. Engl. J. Med.*, **2**, 1171–1175.
3. Jedlitschky, G., Leier, I., Buchholz, U., Hummel-Eisenbeiss, J., Burchell, B. and Keppler, D. (1997) ATP-dependent transport of bilirubin glucuronides by the multidrug resistance protein MRP1 and its hepatocyte canalicular isoform MRP2. *Biochem. J.*, **327**, 305–310.
4. Pilia, G., Chen, W.M., Scuteri, A., Orrú, M., Albai, G., Dei, M., Lai, S., Usala, G., Lai, M., Loi, P. *et al.* (2006) Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet.*, **2**, e132. doi:10.1371/journal.pgen.0020132.
5. Scuteri, A., Sanna, S., Chen, W.M., Uda, M., Albai, G., Strait, J., Najjar, S., Nagaraja, R., Orrú, M., Usala, G. *et al.* (2007) Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet.*, **3**, e115.
6. Van Es, H.H., Bout, A., Liu, J., Anderson, L., Duncan, A.M., Bosma, P., Oude Elferink, R., Jansen, P.L., Chowdhury, J.R. and Schurr, E. (1993) Assignment of the human UDP glucuronosyltransferase gene (UGT1A1) to chromosome region 2q37. *Cytogenet. Cell. Genet.*, **63**, 114–116.
7. König, J., Cui, Y., Nies, A.T. and Keppler, D. (2000) Localization and genomic organization of a new hepatocellular organic anion transporting polypeptide. *J. Biol. Chem.*, **275**, 23161–23168.
8. Galanello, R., Perseu, L., Melis, M.A., Cipollina, L., Barella, S., Giagu, N., Turco, M.P., Maccioni, O. and Cao, A. (1997) Hyperbilirubinemia in heterozygous beta-thalassaemia in related to co-inherited Gilbert's syndrome. *Br. J. Haematol.*, **99**, 433–436.
9. Iolascon, A., Faienza, M.F., Perrotta, S., Meloni, G.F., Ruggiu, G. and del Giudice, E.M. (1999) Gilbert's syndrome and jaundice in

glucose-6-phosphate dehydrogenase deficient neonates. *Haematologica*, **84**, 99–102.
10. Arnaud-Lopez, L., Usala, G., Ceresini, G., Mitchell, B.D., Pilia, M.G., Piras, M.G., Sestu, N., Maschio, A., Busonero, F., Albai, G. *et al.* (2008) Phosphodiesterase 8B gene variants are associated with serum TSH levels and thyroid function. *Am. J. Hum. Genet.*, **82**, 1270–1280.
11. Mitchell, B.D., McArdle, P.F., Shen, H., Rampersaud, E., Pollin, T.I., Bielak, L.F., Jaquish, C., Douglas, J.A., Roy-Gagnon, M.H., Sack, P. *et al.* (2008) The genetic response to short-term interventions affecting cardiovascular function: rationale and design of the Heredity and Phenotype Intervention (HAPI) Heart Study. *Am. Heart J.*, **155**, 823–828.
12. Abe, T., Kakyo, M., Tokui, T., Nakagomi, R., Nishio, T., Nakai, D., Nomura, H., Unno, M., Suzuki, M., Naitoh, T. *et al.* (1999) Identification of a novel gene family encoding human liver-specific organic anion transporter LST-1. *J. Biol. Chem.*, **274**, 17159–17163.
13. Konug, J., Cui, Y., Nies, A.T. and Kepplera, D. (2000) Novel human organic anion transporting polypeptide localized to the basolateral hepatocyte membrane. *Am. J. Physiol. Gastrointest. Liver Physiol.*, **278**, 156–164.
14. Zhang, W., He, Y.J., Gan, Z., Fan, L., Li, Q., Wang, A., Liu, Z.Q., Deng, S., Huang, Y.F., Xu, L.Y. *et al.* (2007) OATP1B1 polymorphism is a major determinant of serum bilirubin level but not associated with rifampicin-mediated bilirubin elevation. *Clin. Exp. Pharmacol. Physiol.*, **34**, 1240–1244.
15. Nozawa, T., Nakajima, M., Tamai, I., Noda, K., Nezu, J., Sai, Y., Tsuji, A. and Yokoi, T. (2002) Genetic polymorphisms of human organic anion transporters OATP-C (SLC21A6) and OATP-B (SLC21A9): allele frequencies in the Japanese population and functional analysis. *J. Pharmacol. Exp. Ther.*, **302**, 804–813.
16. The International HapMap Consortium (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature*, **449**, 851–861.
17. Sanna, S., Jackson, A.U., Nagaraja, R., Willer, C.J., Chen, W.M., Bonnycastle, L.L., Shen, H., Timpson, N., Lettre, G., Usala, G. *et al.* (2008) Association of human height with genetic variants in the GDF5 - BFZB locus: genome-wide scans and replication. *Nat. Genet.*, **40**, 198–203.
18. Cappellini, M.D., Martinez di Montemuros, F., Sampietro, M., Tavazzi, D. and Fiorelli, G. (1999) The interaction between Gilbert's syndrome and G6PD deficiency influences bilirubin levels. *Br. J. Haematol.*, **104**, 928–929.
19. Cui, Y., König, J., Leier, I., Buchholz, U. and Keppler, D. (2001) Hepatic uptake of bilirubin and its conjugates by the human organic anion transporter SLC21A6. *J. Biol. Chem.*, **276**, 9626–9630.
20. Trauner, M., Arrese, M., Soroka, C.J., Ananthanarayanan, M., Koeppel, T.A., Schlosser, S.F., Suchy, F.J., Keppler, D. and Boyer, J.L. (1997) The rat canalicular conjugate export pump (Mrp2) is down-regulated in intrahepatic and obstructive cholestasis. *Gastroenterology*, **113**, 255–264.
21. Origa, R., Galanello, R., Perseu, L., Tavazzi, D., Domenica Cappellini, M., Terenzani, L., Forni, G.L., Quarta, G., Boetti, T. and Piga, A. (2009) Cholelithiasis in thalassemia major. *Eur. J. Haematol.*, **82**, 22–25.
22. Abecasis, G.R., Cherny, S.S., Cookson, W.O. and Cardon, L.R. (2002) Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.*, **30**, 97–101.
23. Lander, E.S. and Green, P. (1987) Construction of multilocus genetic linkage maps in humans. *Proc. Natl. Acad. Sci. USA*, **84**, 2363–2367.
24. Chen, W-M. and Abecasis, G.R. (2007) Family based association tests for genome wide association scans. *Am. J. Hum. Genet.*, **81**, 913–926.