

Genetic and Functional Identification of the Likely Causative Variant for Cholesterol Gallstone Disease at the *ABCG5/8* Lithogenic Locus

Oliver von Kampen,^{1*} Stephan Buch,^{1*} Michael Nothnagel,² Lorena Azocar,³ Hector Molina,³ Mario Brosch,¹ Wiebke Erhart,¹ Witigo von Schönfels,⁴ Jan Egberts,⁴ Marcus Seeger,¹ Alexander Arlt,¹ Tobias Balschun,⁵ Andre Franke,⁵ Markus M. Lerch,⁶ Julia Mayerle,⁶ Wolfgang Kratzer,⁷ Bernhard O. Boehm,⁸ Klaus Huse,⁹ Bodo Schniewind,⁴ Katharina Tiemann,¹⁰ Zhao-Yan Jiang,¹¹ Tian-Quan Han,¹¹ Balraj Mittal,¹² Anshika Srivastava,¹² Mogens Fenger,¹³ Torben Jørgensen,¹⁴ Ramin Schirin-Sokhan,¹⁸ Anke Tönjes,¹⁵ Henning Wittenburg,¹⁶ Michael Stumvoll,¹⁵ Holger Kalthoff,¹⁷ Frank Lammert,¹⁸ Jürgen Tepel,¹⁹ Klaus Puschel,²¹ Thomas Becker,⁴ Stefan Schreiber,¹ Matthias Platzer,⁹ Henry Völzke,²⁰ Michael Krawczak,² Juan Francisco Miquel,³ Clemens Schafmayer,⁴ and Jochen Hampe¹

The sterolin locus (*ABCG5/ABCG8*) confers susceptibility for cholesterol gallstone disease in humans. Both the responsible variant and the molecular mechanism causing an increased incidence of gallstones in these patients have as yet not been identified. Genetic mapping utilized patient samples from Germany (2,808 cases, 2,089 controls), Chile (680 cases, 442 controls), Denmark (366 cases, 766 controls), India (247 cases, 224 controls), and China (280 cases, 244 controls). Analysis of allelic imbalance in complementary DNA (cDNA) samples from human liver (n = 22) was performed using pyrosequencing. Transiently transfected HEK293 cells were used for [³H]-cholesterol export assays, analysis of protein expression, and localization of allelic constructs. Through fine mapping in German and Chilean samples, an ~250 kB disease-associated interval could be defined for this locus. Lack of allelic imbalance or allelic splicing of the *ABCG5* and *ABCG8* transcripts in human liver limited the search to coding single nucleotide polymorphisms. Subsequent mutation detection and genotyping yielded two disease-associated variants: *ABCG5*-R50C ($P = 4.94 \times 10^{-9}$) and *ABCG8*-D19H ($P = 1.74 \times 10^{-10}$) in high

Abbreviations: GFC, gallstone-free controls; GWAS, genome-wide association study; LD, linkage disequilibrium; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; SNP, single nucleotide polymorphism.

From the ¹Department of Internal Medicine I, Hospital Schleswig-Holstein, Kiel, Germany; ²Institute for Medical Informatics and Statistics, Hospital Schleswig-Holstein, Kiel, Germany; ³Department of Gastroenterology, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile; ⁴Department of General and Thoracic Surgery, Hospital Schleswig-Holstein, Kiel, Germany; ⁵Institute for Clinical Molecular Biology, Hospital Schleswig-Holstein, Kiel, Germany; ⁶Department of Internal Medicine A, University of Greifswald, Greifswald, Germany; ⁷Department of Internal Medicine I, University of Ulm Medical Centre, Ulm, Germany; ⁸Division of Endocrinology and Diabetes, University of Ulm Medical Centre, Ulm, Germany; ⁹Leibniz Institute for Age Research, Fritz Lipmann Institute, Jena, Germany; ¹⁰Institute for Hematopathology Hamburg, Hamburg, Germany; ¹¹Department of Surgery, Shanghai Institute of Digestive Surgery, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; ¹²Department of Genetics, Sanjay Gandhi Post Graduate Institute of Medical Sciences, Lucknow, India; ¹³Department of Clinical Biochemistry, Hvidovre University Hospital, Hvidovre, Denmark; ¹⁴Research Centre for Prevention and Health, Glostrup University Hospital, Glostrup, Denmark; ¹⁵University of Leipzig, Department of Medicine, Division of Endocrinology and Diabetes, and IFB Adiposity Diseases, University Hospital Leipzig, Leipzig, Germany; ¹⁶Department of Gastroenterology and Hepatology, University Hospital Leipzig, Leipzig, Germany; ¹⁷Institute for Experimental Cancer Research, University Hospital Schleswig-Holstein, Kiel, Germany; ¹⁸Department of Internal Medicine II, University Hospital Homburg, Germany; ¹⁹Department of Surgery, Hospital Osnabrück, Osnabrück, Germany; ²⁰Institute of Community Medicine, University of Greifswald, Greifswald, Germany; ²¹Department of Family Medicine, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile.

Received May 11, 2012; accepted July 23, 2012.

Supported by the Deutsche Forschungsgemeinschaft (DFG Ha 3091/3-1), the German Ministry of Education and Research (BMBF) and the Federal State of Schleswig-Holstein through the popgen biobank. WK and BOB received grant support from the Federal State of Baden-Wuerttemberg (Centre of Excellence "Metabolic disorders") and the Deutsche Forschungsgemeinschaft (DFG GrK 1041). The Study of Health in Pomerania (SHIP) is conducted by the Research Network of Community Medicine of the University Medicine Greifswald. SHIP is funded by the BMBF and the Ministries of Cultural Affairs and of Social Affairs Ministry of the Federal State of Mecklenburg-West Pomerania. Projects in Chile were funded by Fondo Nacional Científico y Tecnológico (FONDECYT, grant no. 1080325) and the FONDAP Center for Genome Regulation (grant no. 15090007) to J.F.M., by a joint BMBF/Conicyt project (grant no. 170-2009 to J.H. and J.F.M.), by a joint DFG/Conicyt project (444 CHL-1113/15/0-1 to F.L., J.F.M., H.W. and J.H.), and the Alexander von Humboldt Foundation (to J.F.M.).

*These authors contributed equally to this work.

pairwise linkage disequilibrium ($r^2 = 0.95$). [^3H]-cholesterol export assays of allelic constructs harboring these genetic candidate variants demonstrated increased transport activity (3.2-fold, $P = 0.003$) only for the ABCG8-19H variant, which was also superior in nested logistic regression models in German ($P = 0.018$), Chilean ($P = 0.030$), and Chinese ($P = 0.040$) patient samples. **Conclusion:** This variant thus provides a molecular basis for biliary cholesterol hypersecretion as the mechanism for cholesterol gallstone formation, thereby drawing a link between “postgenomic” and “pregenomic” pathophysiological knowledge about this common complex disorder. (HEPATOLOGY 2013;57:2407-2417)

Gallstone disease is a frequent and economically relevant health problem worldwide.¹⁻³ In Western countries the prevalence of cholelithiasis is ~10%-20%.^{1,4,5} Moreover, between 20% and 40% of gallstone patients become symptomatic or develop complications,⁶ and the annual number of cholecystectomies exceeds 700,000 in the U.S. and 170,000 in Germany.^{7,8} Economically, gallstone disease has been identified as the second most costly disorder of the digestive tract.⁸ A genetic component in the susceptibility to cholesterol gallstones has been recognized as early as 1937.^{9,10} There is compelling evidence for familial clustering of the disease, and an increased concordance rate has been observed in monozygotic as compared to dizygotic twins.¹¹

Variation at the ABCG5/8 locus was initially identified as a susceptibility factor for gallstone disease in a genome-wide association study (GWAS), and this association has since been validated in independent studies of German, Romanian, Swedish, Danish, Chilean, and Chinese cases and controls.¹²⁻¹⁷ The findings on ABCG8 are also consistent with an earlier identification, in QTL mapping studies, of its murine ortholog at the mouse lithogenic locus *Lith9*.¹⁸ However, although coding variant ABCG8-D19H has often been used as the major tagging single nucleotide polymorphism (SNP) in both the discovery and the replication studies in humans, its causative role is by no means clear. In fact, identification of the disease-causing variant(s) at ABCG5/8 is still lacking, and no functional studies clarifying the likely disease mechanism have been reported so far.

Here we performed a series of experiments, combining genetic mapping in different populations and functional assays, with a view to identify the causative

mechanisms of gallstone formation at the ABCG5/8 locus (Fig. 1). Our results draw a link between the classical “pregenomic” pathophysiological knowledge of the disease, namely, biliary cholesterol supersaturation, and a disease-associated genetic variant in this common complex disorder.^{19,20}

Materials and Methods

Patient Recruitment and Characterization. All patients with gallstone disease had undergone cholecystectomy or were diagnosed with cholelithiasis using B-mode ultrasonography. The gallstone-free controls (GSF) were confirmed to be gallstone-free by ultrasonography or computed tomography. Details about recruitment and clinical characterization has been reported previously for the German,^{12,21,22} Danish,²³ Chilean,² Indian,²⁴ and Chinese²⁵ patients. All patients and controls gave written informed consent prior to the study, and all study protocols were approved by the Institutional Review Board and Ethics Committees at the respective sites. The overall study protocol was approved by the Ethics Committee of the Kiel Medical Faculty (Ethikkommission der Medizinischen Fakultät der Christian-Albrechts-Universität Kiel, #A156/03).

SNP Selection, Genotyping, and Data Analysis. A set of tagging SNPs covering the extended ABCG8/ABCG5 gene region was selected from the CEU HAPMAP dataset (release 28) using Haploview²⁶ with the following parameters: minor allele frequency $\geq 3\%$, pairwise $r^2 \geq 0.8$ between SNPs, $P > 0.05$ for the HWE test in controls, including some redundant SNPs for more robust coverage. In addition, all coding SNPs, splice site SNPs, 5' untranslated region (UTR)

Address reprint requests to: Jochen Hampe, Department of Internal Medicine I, University Hospital Schleswig-Holstein, Schittenhelmstrasse 12, 24105 Kiel, Germany. E-mail: jhampe@1med.uni-kiel.de; fax: +49-431-597-1842, or Prof. Dr. Juan Francisco Miquel, Department of Gastroenterology, Faculty of Medicine, Pontificia Universidad Católica de Chile, Marcoleta 367, Santiago, Chile. E-mail: jfmiquel@med.puc.cl; fax: +56-2-6397780.

Copyright © 2012 by the American Association for the Study of Liver Diseases.

View this article online at wileyonlinelibrary.com.

DOI 10.1002/hep.26009

Potential conflict of interest: Nothing to report.

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

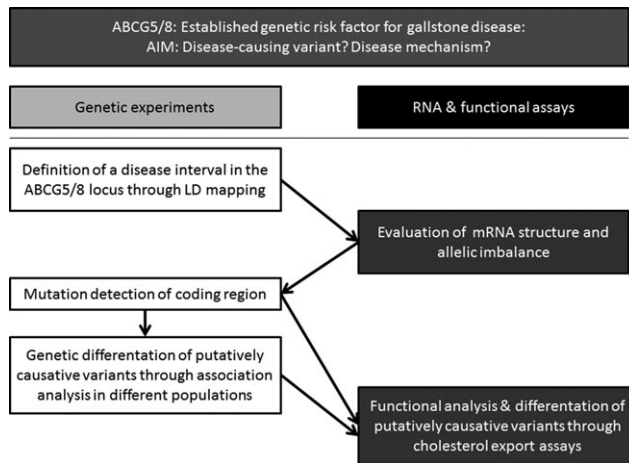


Fig. 1. Study work flow. The scheme gives an overview of the link between genetic and functional experiments performed in the present work.

SNPs, and SNPs located in the intergenic region of *ABCG5* and *ABCG8* that had a minor allele frequency >0.001 in Caucasians, as reported in dbSNP (release 132) or in the scientific literature, or that were detected by Sanger sequencing in our own mutation search were included in the genotyping. Genotyping was performed using either the Sequenom or the Taqman platform as described^{27,28} (see Supporting Methods for further details). All markers were tested for a possible deviation from Hardy-Weinberg equilibrium in the controls before inclusion in the subsequent statistical analysis. Single marker association tests were performed using Haploview²⁶ and PLINK,²⁹ comprising χ^2 or Fisher's exact tests for contingency tables, as appropriate. Logistic regression analysis was performed using R (www.r-project.org) and SPSS (PASW Statistics 18).

Mutation Detection at the *ABCG5* and *ABCG8* Loci. All coding sequence and adjacent splice sites located in the disease-associated region were investigated for potential disease mutations by Sanger sequencing. Primer sequences are provided in the Supporting Methods.

Tissue Samples, Pyrosequencing, and Reverse-Transcription Polymerase Chain Reaction (RT-PCR). Human liver samples were obtained either by surgical or percutaneous biopsy. Total RNA from human liver tissue was isolated using the RNeasy kit from Qiagen (Hilden, Germany) and subsequently reverse-transcribed (Advantage RT-for-PCR kit, Clontech Laboratories, Palo Alto, CA). Matching DNA was isolated from peripheral blood samples. All patients provided written informed consent prior to the study and the sampling protocol was approved by the Ethics

Committee of the Kiel Medical Faculty (Ethikkommission der Medizinischen Fakultät der Christian-Albrechts-Universität Kiel, #D425/07). Pyrosequencing was performed as described³⁰ using primers as reported in the Supporting Methods.

Vectors and Transfection. Expression vectors for wildtype (WT) *ABCG5* and *ABCG8* were constructed by subcloning of the coding sequences of both proteins into plasmids pCEV-Cplus and pEF-HA-neo. The different versions of both transporters, defined by the respective alleles of R50C and D19H, were introduced by site-directed mutagenesis (QuickChange Lightning Site-Directed Mutagenesis Kit, Agilent Technologies, Santa Clara, CA) according to the manufacturer's protocol. HEK cells were transfected using Effectene Transfection Reagent (Qiagen) according to the manufacturer's protocol.

Cholesterol Efflux Assay. Cholesterol efflux was measured as described by Vrins et al.³¹ In brief, HEK cells transiently expressing different allelic variants of *ABCG5*-R50C and *ABCG8*-D19H were incubated in 24-well cell culture plates with 0.5 mL Dulbecco's modified Eagle's medium (DMEM) supplemented with 10 mM HEPES, pH 7.4, 30 mg/mL cholesterol, 0.5 mCi/mL [³H]-cholesterol and 0.2% fatty acid free bovine serum albumin (BSA) for 24 hours. After removal of loading medium, cells were washed four times with DMEM, supplemented with 0.2% fatty acid free BSA, and efflux was initiated by addition of 0.5 mL DMEM supplemented with 10 mM taurocholic acid and 0.5 mM phosphatidylcholine. After 2 hours medium was collected and centrifuged (10,000g, 5 minutes). The remaining cell associated [³H]-cholesterol was determined after extraction for 30 minutes with 0.5 mL isopropanol. Samples of loading medium, efflux medium, and the isopropanol extract were mixed with 4 mL Aquasafe 500 Plus scintillation cocktail (Zinsser Analytic, Frankfurt, Germany) and analyzed with a beta-scintillation counter LS 6500 (Beckman Coulter, Brea, CA). The radioactivity released to the efflux medium was expressed as a fraction of the total radioactive cholesterol. Cholesterol efflux was analyzed compared to the cholesterol efflux of cells coexpressing WT *ABCG5* and *ABCG8* after baseline correction from cells transfected with empty expression vectors only.

Plasma Membrane Detection of *ABCG5*, *G8*, and Allelic Variants and Western Blot Analysis. HEK cells were grown in 10 mL culture dishes for 48 hours to 80% confluency. Biotinylation of cell surface proteins was performed with a Cell Surface Protein Isolation Kit (Pierce, Rockford, IL) according to a modified protocol (see Supporting Methods). Fifteen μ g of each

Table 1. Overview of Patient Samples

Designation	Origin	N _{cases}	Median Age	% Male	BMI	N _{controls}	Median age	% Male	BMI
Mapping	Germany 1	1,266	54	36	27.2	1,000	64	36	26.8
Mapping	Chile 1	167	55	0	29.5	81	55	0	29.9
Differentiation	Germany 2	1,542	59	28	27.4	1,089	61	28	27.6
Differentiation	Chile all	680	55	10	29.1	442	55	10	28.8
Differentiation	Denmark	321	60	38	25.5	816	60	38	26.2
Differentiation	India	247	48	32	na	224	44	40	Na
Differentiation	China	280	54	36	24.6	244	57	36	22.9*

"Mapping samples" were utilized for the definition of the disease interval. "Differentiation" cohorts were genotyped for the assignment of the likely causative mutations among the significant functional variants in the disease interval. The "Chile all" sample included the individuals from the Chilean mapping sample.

*BMI information was only available for 41% of the control individuals.

membrane fraction and total protein lysate were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) with 12% PA-Gels. Proteins were blotted onto 0.45 μ m polyvinylidene difluoride membranes (Millipore, Billerica, MA) and probed with the following primary antibodies from Sigma-Aldrich: monoclonal anti-FLAG M2 antibody, polyclonal anti-HA, and monoclonal β -actin. As secondary antibodies, mouse horseradish peroxidase (HRP)-linked antimouse and antirabbit conjugates (GE Healthcare, Little Chalfont, UK) were used.

Immunofluorescence Microscopy. Cells were grown on 60 mm plates to 80% confluence, transfected as described above with 2 μ g total plasmid complementary DNA (cDNA). Cells were fixed with cold acetone 36 hours after transfection and subsequently permeabilized with 0.1% Triton X-100 for 15 minutes, blocked in 10% goat serum in phosphate-buffered saline (PBS), and incubated overnight with ABCG8 monoclonal antibody (NBP71706, Novus Biologicals, Littleton, CO). After washing with PBS 5 times, cells were incubated with a secondary antimouse antibody (Alexa 568) for 1 hour, washed with PBS 1 \times , and mounted in Fluoromount G (Electron Microscopy Sciences, Hatfield, PA). Labeled sections were examined and captured with a confocal microscope (Olympus BX51, Olympus, Tokyo, Japan).

Results

Linkage Disequilibrium (LD)-Based Definition of the Disease Interval in the ABCG5/8 Locus. The overall study workflow is depicted in Fig. 1. A total of 1,266 German cholelithiasis patients and 1,000 gallstone-free controls ("mapping panel," Table 1) were used to analyze the disease association and the local LD structure using 57 known tagging SNPs at the ABCG5/8 locus (Fig. 2A) selected CEU HAPMAP variants. In a single point analysis (Supporting Table 1),

the most significant disease associations were observed in an allele-based test for variants rs11887534 (ABCG8-D19H, $P = 1.7 \times 10^{-10}$), rs72875462 (ABCG8 intronic, $P = 2.1 \times 10^{-10}$), and rs56132765 (ABCG8-V151V $P = 3.9 \times 10^{-10}$). SNP rs6756629 (ABCG5-R50C) yielded the second most significant association ($P = 4.9 \times 10^{-9}$) of all previously known, nonsynonymous variants in ABCG5/8. The same set of 57 tagging SNPs was also analyzed in a Chilean case control sample, where it yielded a similar pattern of disease association and intermarker LD (Supporting Table 1, Fig. 2B). Based on pairwise $r^2 > 0.5$ to any of the four variants mentioned above, a disease-associated genomic region was defined using the mapping panel, comprising map positions 44.05 Mb to 44.08 Mb (hg19/GRCh37 genome assembly) on chromosome 2. Owing to the opposite orientation of the two ABCG genes, this region contained the five 5' exons of ABCG5 and the six 5' exons of ABCG8.

Disease-Associated Variants Are Not Associated with Allelic Splicing or Allelic Expression. Because formal definition of the disease-associated genomic region on chromosome 2 confined the location of potentially causal variants to the ABCG5 and ABCG8 genes, their transcripts were next evaluated for allelic expression differences. To this end, pyrosequencing assays for exonic variants rs6720173, rs6756629 in ABCG5 and rs11887534, rs56132765, and rs4148217 in ABCG8 were applied to 22 cDNA samples from human liver, including at least four heterozygotes for each variant. No allelic imbalance was detected for any of the five SNPs, including the most significantly disease-associated variant rs11887534 (Fig. 3A). This lack of allelic imbalance was also indicative of an absence of allele-dependent splicing. In order to systematically exclude the presence of variants affecting transcript structure, however, alternative splicing was examined by PCR across both transcripts. To this end, cDNAs from homozygous carriers of the major allele and

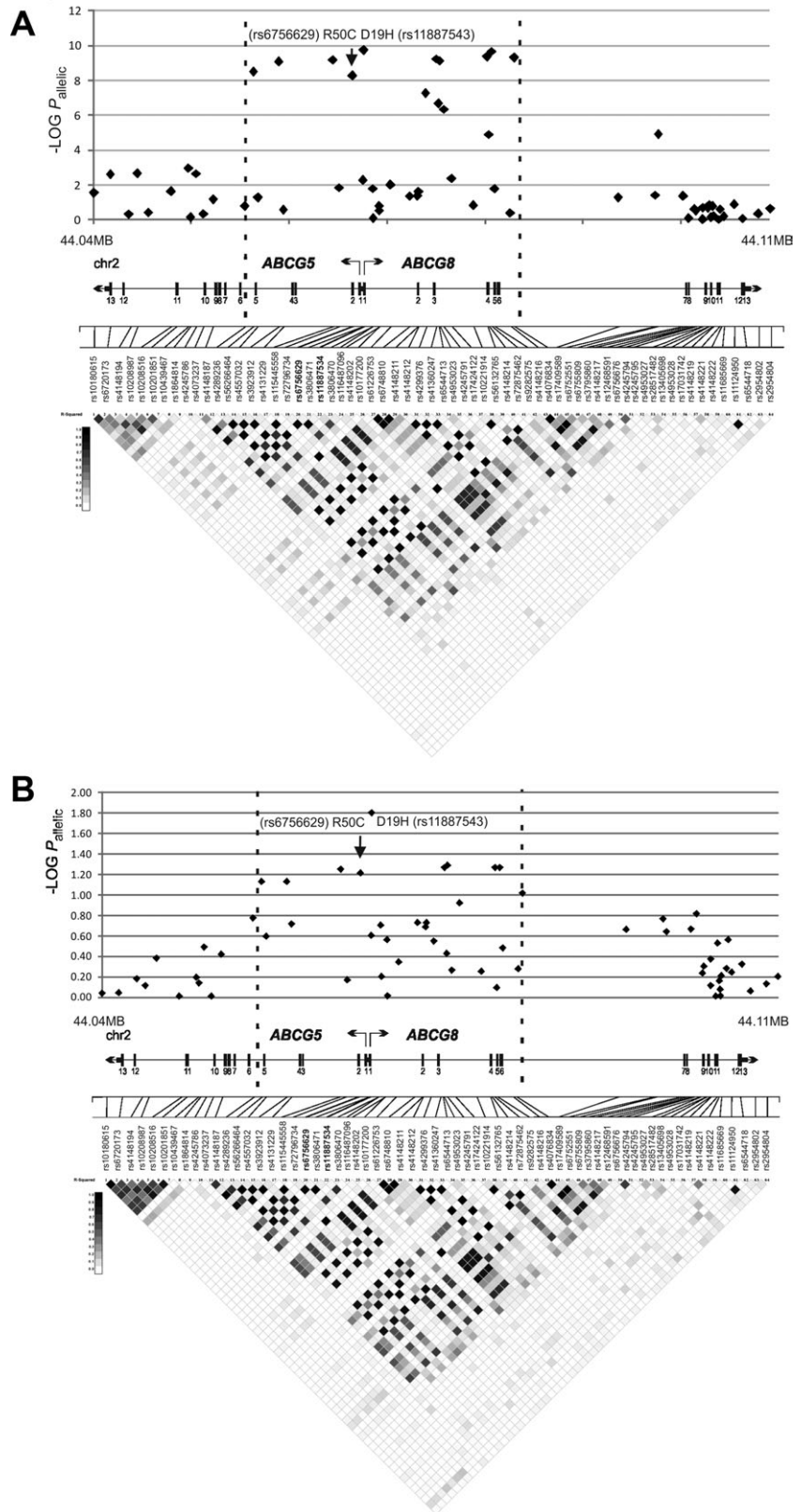


Fig. 2. Definition of the region of disease region in the German (A) and Chilean (B) mapping populations. In each panel the association signal (as the negative Log_{10}) and the pairwise r^2 values are provided in relation to the physical structure of the *ABCG5* and *ABCG8* gene region. The region of disease-associated linkage disequilibrium is indicated by dashed lines in the allelic association plots.

heterozygous individuals for the two most strongly disease-associated variants rs11887534 and rs6756629 were studied in an exon walk. However, no evidence of allele-dependent splicing was found (Fig. 3B).

ABCG8-D19H and ABCG5-R50C Are the Only Functional Candidates in the Disease Interval. Owing to the absence of allelic imbalance and allele-dependent splicing, disease mechanisms such as differential

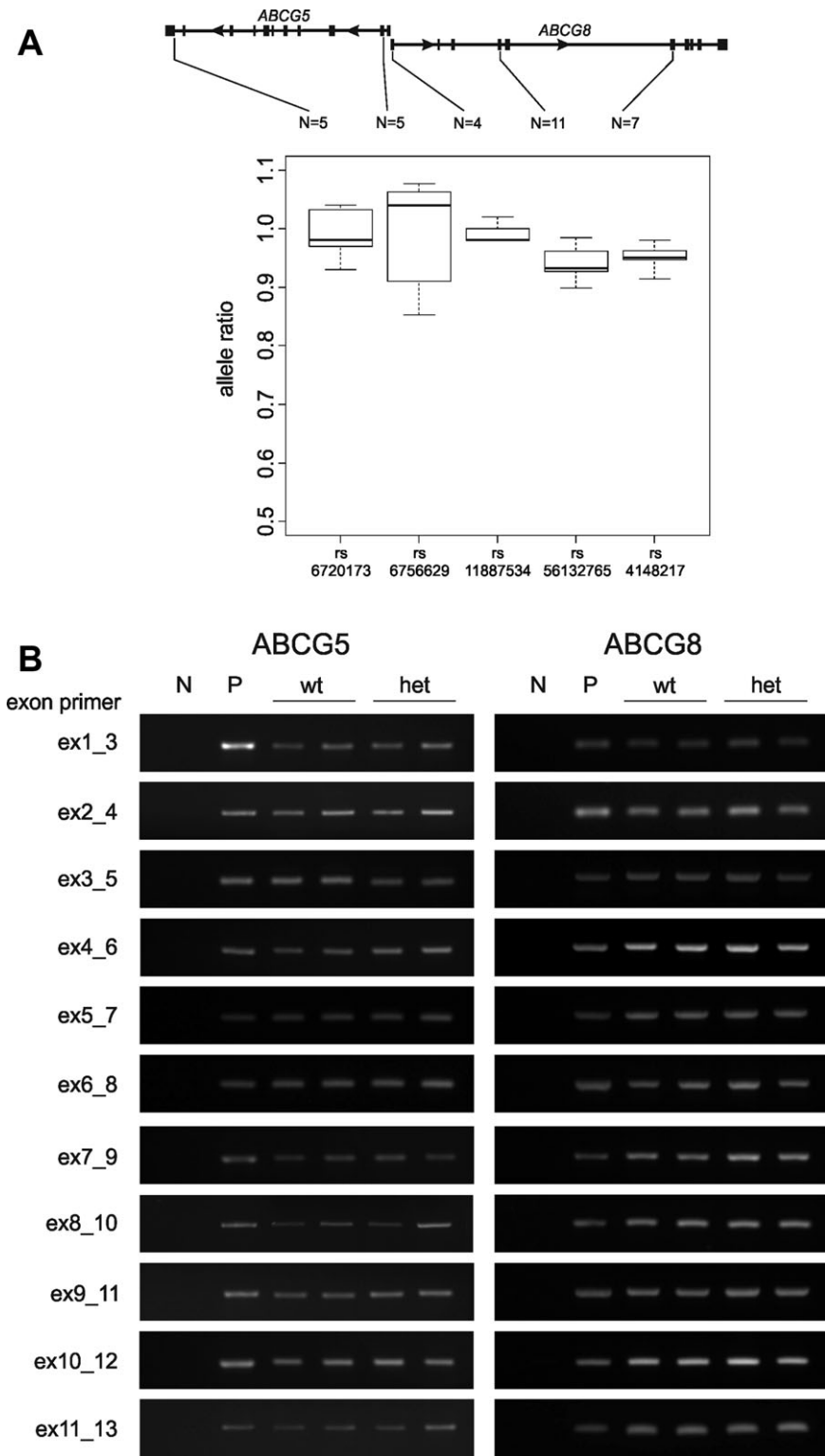


Fig. 3. (A) Evaluation of allelic expression at the *ABCG5* and *ABCG8* transcripts: The transcript ratios as measured by pyrosequencing in liver cDNAs from heterozygote individuals are depicted as boxplots. The number of informative individuals for each variant is provided above the boxplot. The boxplot whiskers indicate the range of the measured allelic ratios. (B) Evaluation of allelic splicing for the *ABCG5* and *ABCG8* transcripts: RT-PCRs in liver cDNA were performed with primer pairs spanning each one exon between the forward and reverse primer (see Supporting Methods). The respective exons are annotated on the left of the figure. Two wildtype (wt) and two heterozygote (het) individuals for both rs11887534 and rs6756629 (the two most strongly associated variants in the disease interval) are shown. N: negative control, P: positive control.

transcription efficiency (caused by promoter polymorphisms or variable intronic enhancers) as well as differential transcript structure or stability could be safely ruled out. Therefore, we next investigated, by Sanger sequencing, all protein-coding DNA sequence in the disease-associated region for potential disease muta-

tions, using samples from 53 cases and 43 controls including 33 carriers of the risk allele of rs11887534 (Supporting Table 2). In addition to known coding variants (according to dbSNP release 132), only one new variant in exon11 of *ABCG8* (Ala548Ala) was detected, namely, in a gallstone patient lacking the risk

allele of rs11887534. All known coding variants (dbSNP release 132) and a selection of synonymous variants from dbSNP were subsequently genotyped in the mapping panel (single point association results in Supporting Table 1). Again, SNP rs11887534 was found to be the most significantly disease associated variant.

ABCG8-D19H, but Not ABCG5-R50C, Leads to Increased Cholesterol Transport. Of all coding SNPs investigated in the disease-associated region, only rs6756629 (ABCG5-R50C; odds ratio [OR]_{allelic} = 1.96, 95% confidence interval [CI]: 1.56-2.47) and rs11887534 (ABCG8-D19H; OR_{allelic} = 2.07, 95% CI: 1.65-2.60) were found to be significantly associated with cholelithiasis in our mapping panel. The significance of their disease associations was approximately seven orders of magnitude higher than that of the next “best” SNP (rs6720173, ABCG5-Q604E, $P_{\text{allelic}} = 0.0024$) which, moreover, is located well outside the disease-associated region defined above. The ORs of the two variants of interest were also found to be very similar and the variants showed high pairwise LD in the German mapping sample ($r^2 = 0.95$), thereby indicating that assessment of their possible causative role by genetic means alone would be difficult.

Because ABCG5 and ABCG8 form a heterodimeric biliary sterol transporter, expression constructs using all four possible allele combinations of the two SNPs were therefore transiently transfected into HEK293 cells. In line with previous observations,³² each transporter constituent was glycosylated only upon coexpression of the other constituent, irrespective of their allelic status, thereby indicating correct shuttling into the cell membrane (Fig. 4; Supporting Fig. 1). In line with these observations, no transporter was detected in the cell membrane fractions when expressed alone. Membrane integration was also confirmed by immunofluorescence assays using antibodies that target ABCG8 (Fig. 5). Thus, expression of ABCG8-19D or ABCG8-19H alone yielded only intracellular fluorescence, whereas coexpression with either ABCG5-50R or ABCG5-50C resulted in a membrane-associated fluorescence signal. Because ABCG5 is retained in the endoplasmic reticulum (ER) unless ABCG8 is expressed in the same cell,³² the membrane associated fluorescence in cells expressing the ABCG5 dimerization partner (i.e., ABCG8-19D or -19H) is an indirect proof of the correct assembly of the two transporter constituents. Signals for ABCG5 and ABCG8 were similar for all four allelic combinations, both in a western blot analysis using monoclonal tag antibodies (Fig. 4) or antibodies directed against human ABCG5 and ABCG8 protein (Supporting Fig. 1), thereby sug-

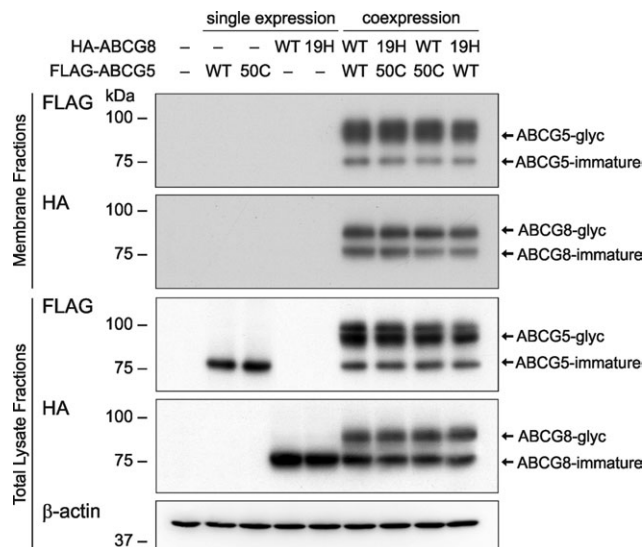


Fig. 4. Western blot analysis of transient expression of allelic constructs of ABC-transporters G5 and G8: HEK cells were transiently transfected with expression constructs allowing expression of FLAG-ABCG5-WT, FLAG-ABCG5-50C, HA-ABCG8-WT, and HA-ABCG8-19H. After 48 hours, total cell lysates and membrane protein fractions (as labeled in the figure) were prepared and subjected to immunoblotting with antibodies targeting the protein tags FLAG and HA. Monoclonal β -actin antibody served as internal loading control. Arrows indicate the unglycosylated (immature) lower and glycosylated (glyc) molecular forms of both ABC-transporters. No evidence of allelic protein expression is seen (see also Supporting Fig. 1). The occurrence of unglycosylated ABC transporter in the membrane fraction upon coexpression with the respective dimerization partner was also seen in a recent publication from other investigators.³¹

gesting that differential protein expression is unlikely to represent the primary disease mechanism.

Next, cholesterol efflux assays using [³H]-labeled cholesterol were performed and revealed a 3.2-fold higher transport efficiency when the ABCG8-19H allele was included in the expression construct (two-sided t test, $P = 0.0032$ and $P = 0.0036$ for the coexpression of either ABCG5-50R or ABCG5-50C, respectively; Fig. 6). Whether ABCG5-50R or ABCG5-50C was present in the construct did not influence cholesterol efflux, given the respective ABCG8-D19H allele. Thus, among the two genetically linked candidate SNPs ($r^2 = 0.95$), we may conclude that an increased sterol transport efficiency mediated by ABCG8-19H alone is a likely functional effect of the two SNPs.

ABCG8-D19H But Not ABCG5-R50C Captures the Genetic Risk Across Populations. In addition to the functional experiments, nested logistic regression analyses were performed to statistically evaluate the causative role of individual variants in the disease-associated region. In the German mapping panel, none of the coding SNPs listed in Supporting Table 1 significantly improved the allelic risk model over the inclusion of ABCG8-D19H alone. When ABCG5-R50C

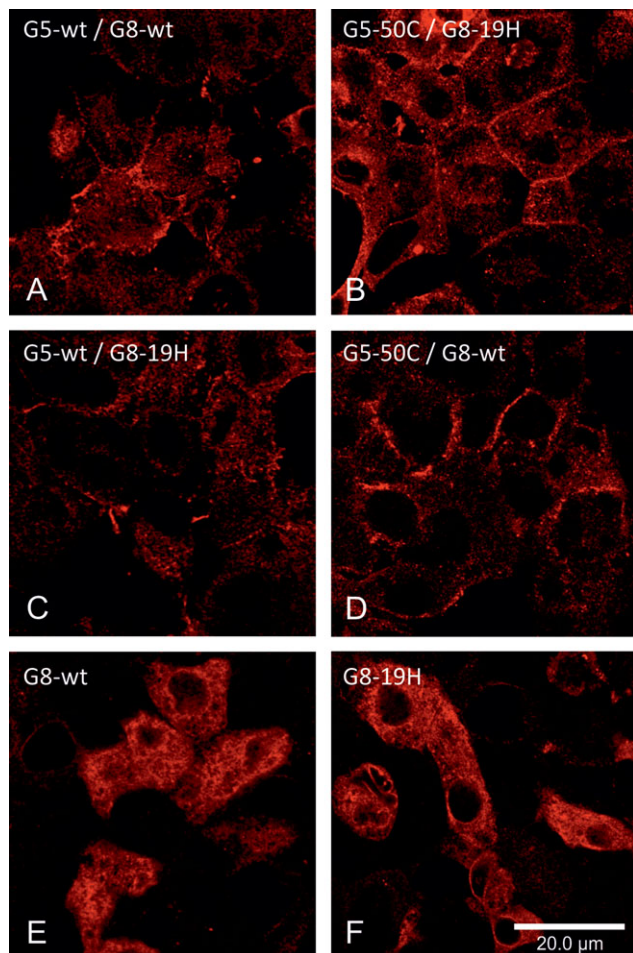


Fig. 5. Subcellular localization and expression of ABCG5, ABCG8, and their allelic variants in immunohistochemistry analysis in HEK cells: The respective allelic combinations are noted in each panel. The red fluorescence signal shows sites of ABCG8 expression (A-F). Membrane-associated fluorescence is found exclusively upon coexpression with ABCG5 or the respective allelic variant (A-D). No allelic differences in expression or subcellular localization are seen. See Supporting Fig. 2 for controls.

was included as a mandatory explanatory variable, ABCG8-D19H significantly improved the model fit ($P = 0.0175$), thereby suggesting again that ABCG8-D19H is the major causative variant in the region. To confirm this result, rs6756629 (ABCG5-R50C) and rs11887534 (ABCG8-D19H) were also genotyped in additional case-controls samples of German, Chilean, Danish, Indian, and Chinese origin (Table 2). Sample size and haplotype structure allowed differentiation between the effects of the two SNPs in the Chilean ($P = 0.030$) and Chinese substudy ($P = 0.040$), in the latter owing to the lower level of LD between the two variants in this population ($r^2 = 0.73$) despite the limited sample size. In each of the samples, ABCG8-D19H consistently showed an equal or higher allelic OR as compared to ABCG5-R50C, thereby support-

ing the above conclusion. Moreover, logistic regression analysis containing ABCG8-D19H and other tagging SNPs (Supporting Table 1) were performed in the German and Chilean mapping panels also revealing no significant improvement over ABCG8-D19H alone (all $P > 0.1$). Thus, a prominent role of hitherto undetected rare variants tagged by one of the other SNPs at the locus of interest is also unlikely.

Discussion

Biliary cholesterol hypersaturation has long been established, through bile analysis in patients and through animal models, as one of the key mechanisms of gallstone formation.³³⁻³⁷ With the introduction of Fourier Transform Infrared Spectroscopy (FTIR) as a means of analyzing gallstone composition in the 1970s,^{38,39} cholesterol has been confirmed as the predominant compound of gallstones.⁴⁰ Indeed, a recent analysis of over 1,000 samples from Germany revealed cholesterol as the most abundant substance in over 93% of gallstones.⁴¹ ABCG5 and ABCG8 form obligate heterodimers for biliary sterol secretion and export

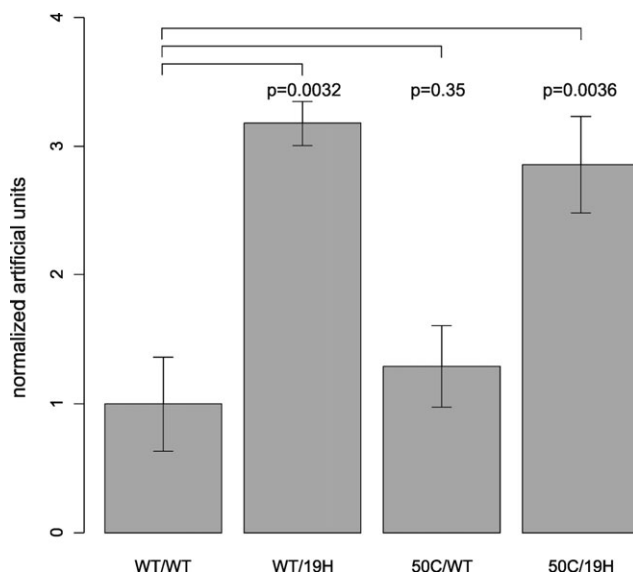


Fig. 6. Cholesterol efflux mediated by ABCG5/G8 and their allelic variants ABCG5-50C and ABCG8-19H: HEK cells transiently coexpressing ABCG5 and ABCG8 and their allelic variants in all four combinations were loaded with [³H]-cholesterol for 24 hours. Subsequently, the medium was replaced against efflux medium and the cells were incubated for 2 hours. Cholesterol efflux was calculated and expressed arbitrary units normalized on the efflux of the wildtype (ABCG5-50R and ABCG8-19D). Data are represented as mean (\pm standard deviation) and experiments were performed in triplicate. The P -values in comparison to the wildtype constructs are indicated with brackets. Supporting Table 5 and Supporting Fig. 5 provide the radioactivity counts at time 0 hours. Counts for time 2 hours are shown in Supporting Table 6.

Table 2. Results of Single Marker Association Tests (Denoted "Single") and Nested Logistic Regression Analysis of ABCG5-R50C and ABCG8-D19H

	Cases N	Controls N	R50C			D19H			Nested Models	
			F _{ca}	F _{co}	OR _{allelic} [CI:95%]	F _{ca}	F _{co}	OR _{allelic} [CI:95%]	R50C D19H	D19H R50C
Germany 1	1,266	1,000	0.104	0.056	1.96 [1.56-2.47]	0.110	0.056	2.07 [1.65-2.60]	0.079	0.018
Germany 2	1,542	1,089	0.102	0.049	2.23 [1.78-2.80]	0.107	0.051	2.23 [1.79-2.79]	0.929	0.248
Denmark	336	766	0.089	0.053	1.73 [1.22-2.45]	0.092	0.055	1.75 [1.24-2.47]	0.405	0.376
Chile all	680	442	0.124	0.075	1.75 [1.29-2.35]	0.123	0.070	1.85 [1.37-2.51]	0.174	0.030
India	247	224	0.047	0.022	2.14 [1.01-4.54]	0.047	0.022	2.14 [1.01-4.54]	0.963	0.963
China	280	244	0.009	0.012	0.72 [0.21-2.38]	0.011	0.008	1.31 [0.37-4.67]	0.052	0.040

F_{ca}: allele frequency in cases, F_{co}: allele frequency in controls, R50C | D19H denotes R50C introduced to the model given D19H is contained in the model. The allelic odds ratios of D19H are consistently equal or greater than for R50C in all patient samples. The *P*-value of patient samples, where successful differentiation of the effect between D19H and R50C was possible is denoted in bold.

cholesterol from the canalicular membrane of hepatocytes into the biliary lumen. Loss of function mutations of ABCG5/8 cause a rare autosomal recessive disorder, known as sitosterolemia, that is characterized by increased fractional absorption and decreased biliary secretion of neutral sterols.⁴² In contrast, hepatic overexpression of ABCG5/8 increases hepatic sterol transport and biliary sterol content.⁴³ In the present study we were able to link a statistically inferred candidate disease mutation at the *ABCG5/8* locus to this pathomechanism. More specifically, we could show that the histidine-encoding allele of rs11887534 (ABCG8-D19H) is associated with an increased allelic cholesterol transport efficiency of the ABCG5/8 heterodimer, thereby contributing to biliary cholesterol hypersecretion and supersaturation and, most likely, to a predisposition to gallstone formation. This increased biliary cholesterol transport may also contribute to the lower serum cholesterol and phytosterol levels observed previously in carriers of ABCG8-D19H.^{44,45}

As a first step towards a functional resolution of the observed disease association, we searched for genetic variants affecting messenger RNA (mRNA) abundance or structure. To this end, we evaluated the degree of allelic imbalance (AI) of both candidate transcripts in human liver—the human target tissue of interest. Compared to overall expression analysis, the assessment of allelic imbalance⁴⁶ inherently provides an internal control, namely, the second copy of the gene. Allele-dependent transcription is therefore detected more specifically using this approach because overall changes in transcription level, particularly in human liver, where the latter is influenced by many factors, including nutrition status, medications, and concomitant disorders, do not influence the AI readout. In any case, the absence of allelic imbalance at all five exonic SNPs located in the two transcripts effectively ruled out variable transcription efficiency, transcript stability, or alternative splicing as possible disease mechanisms.

Further searches for the disease-causing variant(s) could thus be confined to the coding sequence of the two genes. After extensive genotyping, only two candidate variants remained: ABCG5-R50C and ABCG8-D19H, both with similar allelic odds ratios (Supporting Table 1; Table 2). Because the two SNPs were in high LD, formal genetic differentiation of the two variants was achieved only in three out of the six investigated patient samples from Germany, Denmark, Chile, India, and China, implicating rs11887534 (ABCG8-D19H) as the likely risk factor. Differentiation in the Chinese patient sample was aided by the markedly different LD structure in this population. In line with these statistical results, only ABCG8-D19H was found to yield a functional effect in *in vitro* coexpression experiments.

Despite our systematic and intensive mutation search, it cannot be ruled out that hitherto undiscovered rare variants at the *ABCG5/8* locus also play a causative role in cholelithiasis. However, in a logistic regression analysis none of the known common tagging SNPs improved the model fit significantly over rs11887534 (ABCG8-D19H), which implies that there is probably not much "missing heritability" left at this locus.

The role of the ABCG5/8 heterodimer as a biliary cholesterol transporter and cholesterol supersaturation as a key pathomechanism of gallstone formation on the functional side and the association of the *ABCG5/8* locus with gallstone risk using rs11887534 as the tagging variant were known previously. In this report, we genetically establish rs11887534 (ABCG8-D19H) through thorough transcript mapping, mutation detection, and association analysis in ethnically different populations as the likely causative variant for gallstone susceptibility. Further, we show that rs11887534 (ABCG8-D19H) increases the transport efficiency of cholesterol of the sterolin heterodimer, thereby drawing a link between the known basic pathomechanism

of cholesterol gallstone formation to this specific genetic variant.

References

- Kratzer W, Mason RA, Kachele V. Prevalence of gallstones in sonographic surveys worldwide. *J Clin Ultrasound* 1999;27:1-7.
- Miquel JF, Covarrubias C, Villaroel L, Mingrone G, Greco AV, Puglielli L, et al. Genetic epidemiology of cholesterol cholelithiasis among Chilean Hispanics, Amerindians, and Maoris. *Gastroenterology* 1998;115:937-946.
- Portincasa P, Moschetta A, Palasciano G. Cholesterol gallstone disease. *Lancet* 2006;368:230-239.
- Kratzer W, Kachele V, Mason RA, Hill V, Hay B, Haug C, et al. Gallstone prevalence in Germany: the Ulm Gallbladder Stone Study. *Dig Dis Sci* 1998;43:1285-1291.
- Kratzer W, Kron M, Hay B, Pfeiffer MM, Kachele V. [Prevalence of cholecystolithiasis in South Germany—an ultrasound study of 2,498 persons of a rural population.] *Z Gastroenterol* 1999;37:1157-1162.
- Gibney EJ. Asymptomatic gallstones. *Br J Surg* 1990;77:368-372.
- BQS Bundesgeschäftsstelle Qualitätssicherung GmbH. Cholezystektomie. Qualität sichtbar machen — BQS-Qualitätsreport 2004 [Cholezystektomie. Making quality visible — BQS Quality Report 2004]. Düsseldorf, Germany: Bundesgeschäftsstelle Qualitätssicherung, ed. 2008.
- Sandler RS, Everhart JE, Donowitz M, Adams E, Cronin K, Goodman C, et al. The burden of selected digestive diseases in the United States. *Gastroenterology* 2002;122:1500-1511.
- Körner G. Über die familiäre Häufung der Gallenblasenkrankheiten [About the familiar accumulation of gallbladder disease]. *Zeitschrift für Menschliche Vererbungs- und Konstitutionslehre* 1937;20:528-582.
- Nürnberg D, Berndt H, Pannwitz H. Familiäre Häufung von Gallensteinen [Family history of gallstones]. *Dtsch med Wschr* 1989;114:1059-1063.
- Katsika D, Grjibovski A, Einarsson C, Lammert F, Lichtenstein P, Marschall HU. Genetic and environmental influences on symptomatic gallstone disease: a Swedish study of 43,141 twin pairs. *HEPATOLOGY* 2005;41:1138-1143.
- Buch S, Schafmayer C, Volzke H, Becker C, Franke A, von Eller-Eberstein H, et al. A genome-wide association scan identifies the hepatic cholesterol transporter ABCG8 as a susceptibility factor for human gallstone disease. *Nat Genet* 2007;39:995-999.
- Grunhage F, Acalovschi M, Tirziu S, Walier M, Wienker TF, Ciocan A, et al. Increased gallstone risk in humans conferred by common variant of hepatic ATP-binding cassette transporter for cholesterol. *HEPATOLOGY* 2007;46:793-801.
- Katsika D, Magnusson P, Krawczyk M, Grunhage F, Lichtenstein P, Einarsson C, et al. Gallstone disease in Swedish twins: risk is associated with ABCG8 D19H genotype. *J Intern Med*. 2010;268:279-285.
- Kuo KK, Shin SJ, Chen ZC, Yang YH, Yang JF, Hsiao PJ. Significant association of ABCG5 604Q and ABCG8 D19H polymorphisms with gallstone disease. *Br J Surg* 2008;95:1005-1011.
- Marschall HU, Katsika D, Rudling M, Einarsson C. The genetic background of gallstone formation: an update. *Biochem Biophys Res Commun* 2010;396:58-62.
- Stender S, Frikke-Schmidt R, Nordestgaard BG, Tybjaerg-Hansen A. Sterol transporter adenosine triphosphate-binding cassette transporter G8, gallstones, and biliary cancer in 62,000 individuals from the general population. *HEPATOLOGY* 2011;53:640-648.
- Wittenburg H, Lyons MA, Li R, Churchill GA, Carey MC, Paigen B. FXR and ABCG5/ABCG8 as determinants of cholesterol gallstone formation from quantitative trait locus mapping in mice. *Gastroenterology* 2003;125:868-881.
- Van Erpecum KJ. Pathogenesis of cholesterol and pigment gallstones: an update. *Clin Res Hepatol Gastroenterol* 2011;35:281-287.
- Wittenburg H, Lammert F. Genetic predisposition to gallbladder stones. *Semin Liver Dis* 2007;27:109-121.
- Volzke H, Baumeister SE, Alte D, Hoffmann W, Schwahn C, Simon P, et al. Independent risk factors for gallstone formation in a region with high cholelithiasis prevalence. *Digestion* 2005;71:97-105.
- Steffens M, Lamina C, Illig T, Bettecken T, Vogler R, Entz P, et al. SNP-based analysis of genetic substructure in the German population. *Hum Hered* 2006;62:20-29.
- Friedrich N, Volzke H, Hampe J, Jorgensen T. Known risk factors do not explain the differences in gallstone prevalence between Denmark and North East Germany. *Am J Gastroenterol* 2009;104:89-95.
- Jindal C, Kumar S, Choudhari G, Goel H, Mittal B. Organic anion transporter protein (OATP1B1) encoded by SLCO1B1 gene polymorphism (388A>G) and susceptibility in gallstone disease. *Indian J Med Res* 2009;129:170-175.
- Wang Y, Jiang ZY, Fei J, Xin L, Cai Q, Jiang ZH, et al. ATP binding cassette G8 T400K polymorphism may affect the risk of gallstone disease among Chinese males. *Clin Chim Acta* 2007;384:80-85.
- Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263-265.
- Hampe J, Franke A, Rosenstiel P, Till A, Teuber M, Huse K, et al. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 2007;39:207-211.
- Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD, et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 2011;43:246-252.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559-575.
- Utting M, Hampe J, Platzer M, Huse K. Locking of 3' ends of single-stranded DNA templates for improved pyrosequencing performance. *Biotechniques* 2004;37:66-67.
- Vrins C, Vink E, Vandenbergh KE, Frijters R, Seppen J, Groen AK. The sterol transporting heterodimer ABCG5/ABCG8 requires bile salts to mediate cholesterol efflux. *FEBS Lett* 2007;581:4616-4620.
- Graf GA, Li WP, Gerard RD, Gelissen I, White A, Cohen JC, et al. Coexpression of ATP-binding cassette proteins ABCG5 and ABCG8 permits their transport to the apical surface. *J Clin Invest* 2002;110:659-669.
- Gustafsson U, Benthin L, Granstrom L, Groen AK, Sahlin S, Einarsson C. Changes in gallbladder bile composition and crystal detection time in morbidly obese subjects after bariatric surgery. *HEPATOLOGY* 2005;41:1322-1328.
- Venneman NG, Portincasa P, Vanberge-Henegouwen GP, van Erpecum KJ. Cholesterol saturation rather than phospholipid/bile salt ratio or protein content affects crystallization sequences in human gallbladder bile. *Eur J Clin Invest* 2004;34:656-663.
- Carey MC, Small DM. The physical chemistry of cholesterol solubility in bile. Relationship to gallstone formation and dissolution in man. *J Clin Invest* 1978;61:998-1026.
- Admirand WH, Small DM. The physicochemical basis of cholesterol gallstone formation in man. *J Clin Invest* 1968;47:1043-1052.
- Miquel JF, Nunez L, Amigo L, Gonzalez S, Raddatz A, Rigotti A, et al. Cholesterol saturation, not proteins or cholecystitis, is critical for crystal formation in human gallbladder bile. *Gastroenterology* 1998;114:1016-1023.
- Chihara G, Yamamoto S, Kameda H. Medical and biochemical application of infrared absorption spectra. I. Studies on gallstone by infrared spectra and X-ray crystallography. *Chem Pharm Bull (Tokyo)* 1958;6:50-54.
- Trotman BW, Ostrow JD, Soloway RD. Pigment vs cholesterol cholelithiasis: comparison of stone and bile composition. *Am J Dig Dis* 1974;19:585-590.
- Malet PF, Weston NE, Trotman BW, Soloway RD. Cyclic deposition of calcium salts during growth of cholesterol gallstones. *Scan Electron Microsc* 1985;775-779.

41. Schafmayer C, Hartleb J, Tepel J, Albers S, Freitag S, Volzke H, et al. Predictors of gallstone composition in 1025 symptomatic gallstones from Northern Germany. *BMC Gastroenterol* 2006;6:36.
42. Berge KE, Tian H, Graf GA, Yu L, Grishin NV, Schultz J, et al. Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. *Science* 2000;290:1771-1775.
43. Wu JE, Basso F, Shamburek RD, Amar MJ, Vaisman B, Szakacs G, et al. Hepatic ABCG5 and ABCG8 overexpression increases hepatobiliary sterol transport but does not alter aortic atherosclerosis in transgenic mice. *J Biol Chem* 2004;279:22913-22925.
44. Gylling H, Hallikainen M, Pihlajamaki J, Agren J, Laakso M, Rajaratnam RA, et al. Polymorphisms in the ABCG5 and ABCG8 genes associate with cholesterol absorption and insulin sensitivity. *J Lipid Res* 2004;45:1660-1665.
45. Berge KE, von Bergmann K, Lutjohann D, Guerra R, Grundy SM, Hobbs HH, et al. Heritability of plasma noncholesterol sterols and relationship to DNA sequence polymorphism in ABCG5 and ABCG8. *J Lipid Res* 2002;43:486-494.
46. Oetting WS. Identifying functional promoter SNPs using allelic imbalance. *Hum Mutat* 2010;31:v.