Published in final edited form as:
J Hepatol. 2009 October ; 51(4): 750–757. doi:10.1016/j.jhep.2009.04.027.

Multiple variants in toll-like receptor 4 gene modulate risk of liver fibrosis in Caucasians with chronic hepatitis C infection☆

Li Yonghong[1,*], **Monica Chang**[1], **Olivia Abar**[1,#], **Veronica Garcia**[1], **Charles Rowland**[1], **Joseph Catanese**[1], **David Ross**[1], **Samuel Broder**[1], **Mitchell Shiffman**[2], **Ramsey Cheung**[3], **Teresa Wright**[4,‡], **Scott L. Friedman**[5], and **John Sninsky**[1]

[1]Celera Corporation, 1401 Harbor Bay Parkway, Alameda, CA 94502, USA

[2]Virginia Commonwealth University, Richmond, VA, USA

[3]Stanford University, Stanford, CA, USA

[4]University of California, San Francisco, CA, USA

[5]Mount Sinai School of Medicine, Division of Liver Diseases, New York, NY, USA

Abstract

Background/Aims—Seven genomic loci, implicated by single nucleotide polymorphisms (SNPs), have recently been associated with progression to advanced fibrosis (fibrosis risk) in patients with chronic hepatitis C virus. Other variants in these loci have not been examined but may be associated with fibrosis risk independently of or due to linkage disequilibrium with the original polymorphisms.

Methods—We carried out dense genotyping and association testing of additional SNPs in each of the 7 regions in Caucasian case control samples.

Results—We identified several SNPs in the toll-like receptor 4 (*TLR4*) and syntaxin binding protein 5-like (*STXBP5L*) loci that were associated with fibrosis risk independently of the original significant SNPs. Haplotypes consisting of these SNPs in *TLR4* and *STXBP5L* were strongly associated with fibrosis risk (global $P = 3.04 \times 10^{-5}$ and 4.49×10^{-6}, respectively).

Conclusions—Multiple variants in *TLR4* and *STXBP5L* genes modulate risk of liver fibrosis. These findings are of relevance for understanding the pathogenesis of HCV-induced liver disease in Caucasians and may be extended to other ethnicities as well.

Keywords

Liver fibrosis; Risk factor; Single nucleotide polymorphism; Toll-like receptor TLR4; Syntaxin binding protein STXBP5L

☆Y.L., M.C., V.G., C.R., J.C., D.R., S.B. and J.S. are employees of Celera Corporation and declared their financial interest in the company; O.A. was an employee of Celera Corporation at the time the study was carried out; T.W. declared that she received funding from the drug companies involved in order to carry out her research in this manuscript; M.S., R.C. and S.L.F. declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

[*]Corresponding author. Tel.: +1 510 7496283; fax: +1 510 7496200. yonghong.li@celera.com..
[#]Present address: Siemens Healthcare Diagnostics Inc., Berkeley, CA, USA.
[‡]Present address: Roche Molecular Diagnostics, Pleasanton, CA, USA.

Appendix A.

Supplementary data Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jhep.2009.04.027.

1. Introduction

The development of hepatic fibrosis which leads to cirrhosis in patients with chronic hepatitis C virus (HCV) infection results from inflammatory response. This process is associated with marked inter-patient variation which is diffcult to predict [1,2]. The wide spectrum in the rate of fibrosis progression is thought to be modulated by a combination of host genetic factors and other host variables including age, gender, and alcohol intake [3,4]. Recently, a seven gene variant prognostic signature for cirrhosis (CRS7) in patients with chronic hepatitis C has been developed for Caucasian patients and validated in an independent patient cohort with area under the curve (AUC) = 0.73 (95% CI: 0.56–0.89; *P*-value <0.001) [5].

There are strong *a priori* rationale for and functional follow-up study supporting the genes of the prognostic signature, which were initially identified from a large scale case control genetic association study [5]. Among these genes, the toll-like receptor 4 (TLR4), a lipopoly-saccharide-receptor, plays a critical role in pathogen recognition and activation of innate and adaptive immunity [6]. Interaction of HCV and TLR4 signaling is robust although complex: HCV infection can directly induce TLR4 expression [7] and interfere with TLR4 signaling in immune cells [8], and TLR4 signaling itself may regulate HCV replication [9]. In hepatic stellate cells, activation of TLR4 results in the down-regulation of the transforming growth factor (TGF)-β pseudoreceptor Bambi, thereby sensitizing cells to TGF-β-induced signaling leading to hepatic inflammation and fibrosis [10]. Furthermore, a mechanistic study initiated because of these results has already demonstrated that the two fibrosis-associated TLR4 missense variants, T339I (rs4986791) and D299G (rs4986790), have a significant impact on the activity of TLR4 in inflammatory and fibrogenic signaling; more specifically disease protective variants lower the apoptotic threshold of hepatic stellate cells [11]. Of the other genes, the antizyme inhibitor *AZIN1*, considered a tumor suppressor, plays a role in cell proliferation and death [12]. Syntaxin binding protein STXBP5L is likely to be involved in vesicle traf-ficking and exocytosis [13] and may therefore be involved in the replication of HCV in the liver and indirectly in liver fibrosis via promoting an environment conducive to HCV replication.

In this study, we have carried out additional, dense association testing of the above gene regions, for the following reasons. First, we aim to identify likely causal genes or regulatory elements which may be in linkage disequilibrium (LD: the non-random association of alleles at two or more loci) with the original markers. Second, genetic studies have demonstrated high probability for the existence of other, independent risk variants that impair expression and/or function of genes associated with disease risk (see, for example, [14]). Third, if there are independent risk variants at the same locus [15], then this allelic heterogeneity will likely make an important contribution to the phenotype, i.e. disease risk. In particular, the original *TLR4* and *STXBP5L* variants associated with liver fibrosis are rare or absent in Asians and/or Africans [16,17]; and hence identification of additional risk variants at these loci may explain a portion of the disease risk in those populations. It is conceivable that variants other than those reported may modulate disease risk in these as well as Caucasian populations.

2. Materials and methods

2.1. Study design

The individual SNPs of the CRS7 signature were initially identified from a gene-centric, genome-wide associations study of ~25,000 SNPs [5]. Additional SNPs were tested in this follow-up study to provide better coverage of each region implicated by the signature SNPs so that other potentially causal or independently significant markers could be identified. The extent of fine-mapping regions was determined by examining the LD pattern in the HapMap CEPH (Centre d'Etude du Polymorphisme Humain) dataset (www.hapmap.org); we primarily

targeted markers that are present in the same LD region ("main region") as the individual CRS7 markers, although some markers in the adjacent regions were also tested. Markers tested included tagging SNPs (representative SNPs in a region of the genome with high LD), putative functional SNPs, and others such as those in high LD with the individual CRS7 markers. Markers capable of tagging SNP diversity in the main block were selected with the tagger program (http://www.broad.mit.edu/mpg/tagger/server.html) under the following criteria: minor allele frequency ≥0.05 and $r^2 > 0.8$; our sample set had 80% power to detect a variant of 0.05 frequency that has an effect size of 2.2 at the allelic level. The putative functional markers, such as non-synonymous SNPs and those in putative transcription factor binding sites, were selected based on both public and Celera annotation. Additional information for the selected SNPs can be found in Table 1.

2.2. Study samples

The 420 Caucasian samples used in this study were collected from the University of California at San Francisco (UCSF) (N = 187) and the Virginia Commonwealth University (VCU) (N = 233). They consisted of 263 cases and 157 controls where patients with fibrosis stages 3 or 4 were defined as cases and those with fibrosis stage 0 were used as controls; samples with fibrosis stages 1 or 2 were excluded from the study to more effectively delineate genetic factors involved in progression. Fibrosis stages were determined by biopsies read by liver pathologists; the Batts–Ludwig scoring system was utilized in UCSF and the Knodell system in VCU [18]. Cases were sampled at the age of 27–71 years (means ± SD = 49.4 ± 7.4), consisted of 75.3% males, and had daily alcohol intake of 46.2 ± 67.2 g. Controls were sampled at the age of 19–80 years (means ± SD = 47.3 ± 9.1), consisted of 61.1% males, and had daily alcohol intake of 50.5 ± 76.2 g. Sample specific information, including estimated age of infection and duration of infection, is presented in Supplementary Table 1. All patients provided written informed consents, and the study was approved by institutional review boards of UCSF and VCU.

2.3. Genotyping

Cases and controls were individually genotyped by allele-specific, kinetic PCR [19]. For each allele-specific PCR reaction, 0.3 ng of DNA was amplified. Genotypes were automatically called by an in-house software program followed by manual curation without knowledge of case/control status. Our genotyping accuracy was approximately 99% [20].

2.4. Statistical analysis

Allelic association of the SNPs with fibrosis risk was determined by the χ^2 test. Logistic regression was carried out to correct the genetic association for age, gender, alcohol intake and sample source. Logistic regression models for each possible pair of SNPs assumed an additive effect of each additional risk allele on the log odds of fibrosis risk. Linkage disequilibrium (r^2) were calculated from the unphased genotype data using LDMax in the GOLD package [21]. Haplotypes were estimated and tested for association with disease status using a score test with haplotypes coded in an additive fashion [22]. Global tests of association, which test the null hypothesis that the frequency distribution of haplotypes is equal in cases and controls, as well as haplotype specific tests of association were performed [22].

3. Results

3.1. Fine mapping coverage

The individual CRS7 predictor SNPs reside in 7 distinct chromosomal regions (Table 1), where LD extends from ~20 to ~662 kbp according to the HapMap CEPH dataset (www.hapmap.org). To thoroughly examine whether other SNPs in these regions associate with cirrhosis risk more strongly than and/or independently of the original markers, we carried

out dense SNP genotyping in the Caucasian samples used to build the CRS signature [5]. Common HapMap SNPs (of ≥5% allele frequency) in these regions could be effciently tested with a minimum of 14–34 SNPs that tagged other untested markers at $r^2 \geq 0.8$; for the CRS7 predictor 6 region on chromosome 3, marker-marker LD was extensive (~662 kbp), but we only targeted a 163 kbp region that contained *STXBP5* and *POLQ* genes, the only two within this entire LD region, as we were primarily interested in determining which of these genes was more likely to be involved in fibrosis risk. We genotyped a total of 23–71 SNPs for each region; these included tagging, putative functional and other SNPs such as those in high LD with the original marker (Table 1). Coverage of the tagging SNPs by the HapMap markers we tested ranged from 64% to 92% but was likely to be higher since additional non-HapMap markers were genotyped as well.

3.2. The TLR4 locus

We first present the detailed analysis for the *TLR4* locus implicated by the original CRS7 predictor rs4986791. Extensive marker-marker LD was discernable across a region of ~76 kbp encompassing rs4986791 (Fig. 1A). No other genes were located in this region. For fine mapping, we tested an additional 61 SNPs and identified 15 that were associated with cirrhosis risk at allelic $P < 0.05$ (Fig. 1B); the original marker had the strongest effect (Table 2).

Pair-wise SNP regression analysis revealed that significance of some markers could be adjusted away by other significant markers (data not shown), suggesting that all were not independently associated with disease risk, as expected from marker-marker LD (data not shown). Attempting to derive a most parsimonious set of independently significant markers, we identified three groups of SNPs in the *TLR4* region that were associated with fibrosis risk (Table 2). Group 1 contained 9 SNPs including the original *TLR4* marker rs4986791 and 8 other fine mapping SNPs that were in relatively high LD with rs4986791. None of the 8 markers remained significant after adjustment for rs4986791, nor did rs4986791 after adjustment for any of the 8 markers. Of the other significant markers in the *TLR4* region, 5 survived adjustment for rs4986791 (regression $P < 0.05$), four of which were in relatively moderate to high LD (Group 2). An intergenic SNP, rs960312, had the strongest effect and was not independent from the other 3 SNPs in Group 2 as their significance could be adjusted away by each other. The third group contained only one marker, rs11536889, which shared little LD with Group 1 or 2 markers. This marker trended to significance after adjustment for Group 2 marker rs960312 ($P = 0.086$), while rs960312 remained significant after adjustment for the Group 3 marker. Because there was almost no LD between markers in these two groups and they were present in distinct haplotypes (see next), we considered Groups 2 and 3 markers to be independently associated with disease risk.

In multivariant analysis that controls for sample source and other known risk factors including age, gender and alcohol intake, both rs4986791 in Group 1 and rs960312 in Group 2 remained significant at $P < 0.05$ (adjusted $P = 0.0033$ and 0.033, respectively) while rs11536889 in Group 3 trended to significance (adjusted $P = 0.081$). Haplotype analyses with these 3 SNPs resulted in the identification of three significant common haplotypes, each distinguishable by one of the three independent markers (Table 3). The global test of haplotype association was highly significant (global $P = 3.04 \times 10^{-5}$).

3.3. The STXBP5L locus

Of the 58 SNPs tested in the *STXBP5L* locus, 27 were associated with fibrosis risk at allelic $P < 0.05$ (Supplementary Fig. 1). In pair-wise marker regression analysis, two SNPs, rs17740066 and rs2169302, remained significant after adjustment for any of the other markers (Table 4 and data not shown), indicating that no other marker could account for association of these two markers. Conversely, when additional markers were adjusted for rs17740066, only

three remained significant. One of them was rs2169302, and the other two were rs13086038 and rs35827958; the latter two were nearly perfectly concordant ($r^2 = 0.97$). Similarly, a few other markers remained significant when adjusted for rs2169302 but they could be accounted for by rs17740066. Thus the most parsimonious set of independently significant markers at this locus included rs17740066, rs2169302 and rs13086038/rs35827958, all of which remained significant at the level of 0.05 when adjusted for sample source, age, gender and alcohol intake (adjusted $P = 0.00049$, 0.0016 and 0.011/0.011, respectively). LD between these independent markers was low (Table 4; $r^2 < 0.02$ between any pairs). These markers were present in distinct haplotypes (Table 5), and overall haplotype-disease association was strong (global $P = 4.49 \times 10^{-6}$).

3.4. Other loci

The remaining five chromosomal regions were similarly analyzed as above. Although a number of fine mapping markers at each locus were associated with fibrosis risk (Supplementary Fig. 1), none was independent of the original CRS7 predictors (Supplementary Table 2). In the case of SNP predictor 5, association of the original marker rs4290029 could not be accounted for by any other fine mapping marker and all fine mapping markers could be accounted for by rs4290029 (data not shown), suggesting that rs4290029 was likely to be the most informative marker (Supplementary Table 2); SNP5, rs4290029, was located in the intergenic region between *DEGS1* encoding a lipid desaturase and *NVL* encoding a nuclear VCP-like protein. For other loci, relationship between the initial CRS7 markers and other similarly significant and high LD markers we tested could not be teased apart with our regression analysis (Supplementary Table 2). However, these associated markers indicated that *AZIN1* (antizyme inhibitor 1), *TRPM5* (transient receptor potential cation channel, subfamily M, member 5), *AP3S2* (adaptor-related protein complex 3, sigma 2 subunit), and *AQP2* (aquaporin 2) were likely candidate genes modulating liver fibrosis (Supplementary Fig. 1).

In addition to the above analyses, we also tested markers in the regions adjacent to the main LD region that contained the individual CRS7 markers but did not find any other markers that were associated with fibrosis risk independently of the original CRS7 markers (data not shown). This finding was consistent with the fact that markers outside the main LD region shared little to low LD with the original CRS7 markers.

4. Discussion

Our study shows that multiple SNPs in each of the seven chromosomal loci we investigated are significantly associated with an increased risk for developing advanced liver fibrosis and cirrhosis and that this increased risk is specific to certain allele profiles and loci within genes. For the *TLR4* and *STXBP5L* loci, each has three independently significant sets of SNPs, which together give rise to highly significant haplotypes modulating the risk of developing progressive liver fibrosis. For the other 5 loci, each appears to have only one set of independent markers, one of which, rs4290029, an intergenic variant on chromosome 1, may be most informative. Other original and fine mapping markers cannot be distinguished because their association with fibrosis risk can be accounted for by each other. These observations do not exclude the possibility that other less common, independently associated variants may also exist.

The most notable finding of this study is perhaps the *TLR4* locus, which is of considerable interest for its role in immunity as well as genetics since it appears to have undergone selective pressure exerted by pathogens throughout primate evolution [23–25]. To our knowledge, the fine mapping performed for this study represents the most comprehensive analysis of this region, with evaluation of at least 88% of the known SNP diversity covering the entire *TLR4* gene (~13 kbp) and its flanking regulatory sequences. In contrast, previous studies have often

exclusively examined the two co-segregating missense variants T339I (rs4986791) and D299G (rs4986790) in Group 1 (Table 2), both of which are known to attenuate receptor signaling, NFκB activation and pro-inflammatory cytokine production and to impact cell growth and survival [11,26] but they are absent in Asian populations. However, both Group 2 SNP rs960312 and Group 3 SNP rs11536889 are common in Asians (allelic frequency of ~25% in the Hap-Map) and may thus play a role in modulating disease risk in the Asian population. Previously, a putative 3′UTR SNP 11381G/C in *TLR4* has been reported to marginally affect prostate cancer risk ($P = 0.02$) [27]. These non-coding variants may affect gene expression, which may account for a large fraction of disease risk caused by genetic factors [28].

The *TLR4* missense variants appear to be associated with a risk of numerous other diseases and indications as well [29,30], including endotoxin hypo-responsiveness [26], bacterial and fungal infections [31,32], septic shock [33], malaria [34], inflammatory bowel disease [35], atherosclerosis [36], and gastric cancer [37]. Thus, to better assess the overall contribution of *TLR4* polymorphisms on the etiology of these other diseases, the independent variants in Groups 2 and 3 as well as their haplotypes should be further evaluated. In addition, the large number of antagonists and agonists of TLR4 currently being evaluated in drug development strongly suggests that *TLR4* genotypes should be utilized as probable biomarkers in future clinical trials [38,39].

The statistical evidence we presented here is strong, although not at the level reported in some other genome-wide association studies. This is, however, not unexpected, as the number of samples having this type of fibrosis data are limited in contrast to other genome-wide association studies involving tens of thousands of samples. Furthermore, multiple, independent polymorphisms associated in the same gene region support an allelic heterogeneity model of liver fibrosis. While some fine mapping markers (e.g. those in *AZIN1*) and both the *TLR4* and *STXBP5L* haplotypes will remain significant if corrected by the number of tests done in each region, replication in other sample sets of similar characteristics would be fruitful, as for other reported genetic variants [40–47]. The contribution of these additional SNPs determined by regional fine mapping to the CRS merits further investigation, which may be particularly relevant to studies in the Asian and African populations; the previously reported gene variants in *TLR4* and *STXBP5L* are rare in Asians and/or Africans, whereas the gene variants reported here are found at higher frequencies in these populations.

In conclusion we have expanded the list of SNPs and localized putative candidate genes/variants that are independently associated with the risk of liver fibrosis progression and the development of cirrhosis involved in this process. Further examination of these gene variants in mechanistic studies is warranted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to all patients for their participation in this study. We thank clinical staff at the participating university hospitals and Hongjin Huang and colleagues at Celera for excellent technical assistance, Thomas J. White and Andrew Grupe for stimulating discussions, and Steve Schrodi for helpful comments on the manuscript.

Glossary

SNPs	single nucleotide polymorphisms
HCV	hepatitis C virus

TLR4	toll-like receptor 4
STXBP5L	syntaxin binding protein 5-like
AUC	area under the curve
LD	linkage disequilibrium
CEPH	Centre d'Etude du Polymorphisme Humain
UCSF	University of California at San Francisco
VCU	Virginia Commonwealth University

References

[1]. Bataller R, Brenner DA. Liver fibrosis. J Clin Invest 2005;115:209–218. [PubMed: 15690074]

[2]. Friedman SL. Hepatic fibrosis – overview. Toxicology 2008;30:120–129. [PubMed: 18662740]

[3]. Bataller R, North KE, Brenner DA. Genetic polymorphisms and the progression of liver fibrosis: a critical appraisal. Hepatology 2003;37:493–503. [PubMed: 12601343]

[4]. Mallat A, Hezode C, Lotersztajn S. Environmental factors as disease accelerators during chronic hepatitis C. J Hepatol 2008;48:657–665. [PubMed: 18279998]

[5]. Huang H, Shiffman ML, Friedman S, Venkatesh R, Bzowej N, Abar OT, et al. A 7 gene signature identifies the risk of developing cirrhosis in patients with chronic hepatitis C. Hepatology 2007;46:297–306. [PubMed: 17461418]

[6]. Miyake K. Innate immune sensing of pathogens and danger signals by cell surface Toll-like receptors. Semin Immunol 2007;19:3–10. [PubMed: 17275324]

[7]. Machida K, Cheng KT, Sung VM, Levine AM, Foung S, Lai MM. Hepatitis C virus induces toll-like receptor 4 expression, leading to enhanced production of beta interferon and interleukin-6. J Virol 2006;80:866–874. [PubMed: 16378988]

[8]. Agaugue S, Perrin-Cocon L, Andre P, Lotteau V. Hepatitis C lipo-Viro-particle from chronically infected patients interferes with *TLR4* signaling in dendritic cell. PLoS ONE 2007;2:e330. [PubMed: 17389921]

[9]. Broering R, Wu J, Meng Z, Hilgard P, Lu M, Trippler M, et al. Toll-like receptor-stimulated non-parenchymal liver cells can regulate hepatitis C virus replication. J Hepatol 2008;48:914–922. [PubMed: 18362039]

[10]. Seki E, De Minicis S, Osterreicher CH, Kluwe J, Osawa Y, Brenner DA, et al. *TLR4* enhances TGF-beta signaling and hepatic fibrosis. Nat Med 2007;13:1324–1332. [PubMed: 17952090]

[11]. Guo J, Loke J, Zheng F, Yea S, Fugita M, Tarocchi M, et al. Functional linkage of cirrhosis-predictive single nucleotide polymorphisms of Toll-like receptor 4 to hepatic stellate cell response. Hepatology 2009;49:960–968. [PubMed: 19085953]

[12]. Mangold U. Antizyme inhibitor: mysterious modulator of cell proliferation. Cell Mol Life Sci 2006;63:2095–2101. [PubMed: 16847581]

[13]. Yizhar O, Matti U, Melamed R, Hagalili Y, Bruns D, Rettig J, et al. Tomosyn inhibits priming of large dense-core vesicles in a calcium-dependent manner. Proc Natl Acad Sci USA 2004;101:2578–2583. [PubMed: 14983051]

[14]. Li M, Atmaca-Sonmez P, Othman M, Branham KE, Khanna R, Wade MS, et al. CFH haplotypes without the Y402H coding variant show strong association with susceptibility to age-related macular degeneration. Nat Genet 2006;38:1049–1054. [PubMed: 16936733]

[15]. Pritchard JK, Cox NJ. The allelic architecture of human disease genes: common disease-common variant or not? Hum Mol Genet 2002;11:2417–2423. [PubMed: 12351577]

[16]. Hang J, Zhou W, Zhang H, Sun B, Dai H, Su L, et al. TLR4 Asp299Gly and Thr399Ile polymorphisms are very rare in the Chinese population. J Endotoxin Res 2004;10:238–240. [PubMed: 15373967]

[17]. Yoon HJ, Choi JY, Kim CO, Park YS, Kim MS, Kim YK, et al. Lack of Toll-like receptor 4 and 2 polymorphisms in Korean patients with bacteremia. J Korean Med Sci 2006;21:979–982. [PubMed: 17179672]

[18]. Brunt EM. Grading and staging the histopathological lesions of chronic hepatitis: the Knodell histology activity index and beyond. Hepatology 2000;31:241–246. [PubMed: 10613753]

[19]. Germer S, Holland MJ, Higuchi R. High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR. Genome Res 2000;10:258–266. [PubMed: 10673283]

[20]. Li Y, Nowotny P, Holmans P, Smemo S, Kauwe JS, Hinrichs AL, et al. Association of late-onset Alzheimer's disease with genetic variation in multiple members of the GAPD gene family. Proc Natl Acad Sci USA 2004;101:15688–15693. [PubMed: 15507493]

[21]. Abecasis GR, Cookson WO. GOLD – graphical overview of linkage disequilibrium. Bioinformatics 2000;16:182–183. [PubMed: 10842743]

[22]. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 2002;70:425–434. [PubMed: 11791212]

[23]. Smirnova I, Poltorak A, Chan EK, McBride C, Beutler B. Phylogenetic variation and polymorphism at the toll-like receptor 4 locus (TLR4). Genome Biol 2000;1 RESEARCH002.

[24]. Ferwerda B, McCall MB, Alonso S, Giamarellos-Bourboulis EJ, Mouktaroudi M, Izagirre N, et al. TLR4 polymorphisms, infectious diseases, and evolutionary pressure during migration of modern humans. Proc Natl Acad Sci USA 2007;104:16645–16650. [PubMed: 17925445]

[25]. Nakajima T, Ohtani H, Satta Y, Uno Y, Akari H, Ishida T, et al. Natural selection in the TLR-related genes in the course of primate evolution. Immunogenetics 2008;60:727–735. [PubMed: 18810425]

[26]. Arbour NC, Lorenz E, Schutte BC, Zabner J, Kline JN, Jones M, et al. TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. Nat Genet 2000;25:187–191. [PubMed: 10835634]

[27]. Zheng SL, Augustsson-Balter K, Chang B, Hedelin M, Li L, Adami HO, et al. Sequence variants of toll-like receptor 4 are associated with prostate cancer risk: results from the CAncer Prostate in Sweden Study. Cancer Res 2004;64:2918–2922. [PubMed: 15087412]

[28]. Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, et al. Genetics of gene expression and its effect on disease. Nature 2008;452:423–428. [PubMed: 18344981]

[29]. Schroder NW, Schumann RR. Single nucleotide polymorphisms of Toll-like receptors and susceptibility to infectious disease. Lancet Infect Dis 2005;5:156–164. [PubMed: 15766650]

[30]. Ferwerda B, McCall MB, Verheijen K, Kullberg BJ, van der Ven AJ, Van der Meer JW, et al. Functional consequences of toll-like receptor 4 polymorphisms. Mol Med 2008;14:346–352. [PubMed: 18231573]

[31]. Agnese DM, Calvano JE, Hahm SJ, Coyle SM, Corbett SA, Calvano SE, et al. Human toll-like receptor 4 mutations but not CD14 polymorphisms are associated with an increased risk of gram-negative infections. J Infect Dis 2002;186:1522–1525. [PubMed: 12404174]

[32]. Bochud PY, Chien JW, Marr KA, Leisenring WM, Upton A, Janer M, et al. Toll-like receptor 4 polymorphisms and aspergillosis in stem-cell transplantation. N Engl J Med 2008;359:1766–1777. [PubMed: 18946062]

[33]. Lorenz E, Mira JP, Frees KL, Schwartz DA. Relevance of mutations in the TLR4 receptor in patients with gram-negative septic shock. Arch Intern Med 2002;162:1028–1032. [PubMed: 11996613]

[34]. Mockenhaupt FP, Cramer JP, Hamann L, Stegemann MS, Eckert J, Oh NR, et al. Toll-like receptor (TLR) polymorphisms in African children: common TLR-4 variants predispose to severe malaria. Proc Natl Acad Sci USA 2006;103:177–182. [PubMed: 16371473]

[35]. Browning BL, Huebner C, Petermann I, Gearry RB, Barclay ML, Shelling AN, et al. Has toll-like receptor 4 been prematurely dismissed as an inflammatory bowel disease gene? Association study combined with meta-analysis shows strong evidence for association. Am J Gastroenterol 2007;102:2504–2512. [PubMed: 17850411]

[36]. Kiechl S, Lorenz E, Reindl M, Wiedermann CJ, Oberhollenzer F, Bonora E, et al. Toll-like receptor 4 polymorphisms and athero-genesis. N Engl J Med 2002;347:185–192. [PubMed: 12124407]

[37]. Hold GL, Rabkin CS, Chow WH, Smith MG, Gammon MD, Risch HA, et al. A functional polymorphism of toll-like receptor 4 gene increases risk of gastric carcinoma and its precursors. Gastroenterology 2007;32:905–912. [PubMed: 17324405]

[38]. McHutchison J, Poynard T, Afdhal N. Fibrosis as an end point for clinical trials in liver disease: a report of the international fibrosis group. Clin Gastroenterol Hepatol 2006;4:1214–1220. [PubMed: 16979947]

[39]. Kanzler H, Barrat FJ, Hessel EM, Coffman RL. Therapeutic targeting of innate immunity with Toll-like receptor agonists and antagonists. Nat Med 2007;13:552–559. [PubMed: 17479101]

[40]. Halangk J, Sarrazin C, Neumann K, Puhl G, Mueller T, Teuber G, et al. Evaluation of complement factor 5 variants as genetic risk factors for the development of advanced fibrosis in chronic hepatitis C infection. J Hepatol 2008;49:339–345. [PubMed: 18644651]

[41]. Hillebrandt S, Wasmuth HE, Weiskirchen R, Hellerbrand C, Keppeler H, Werth A, et al. Complement factor 5 is a quantitative trait gene that modifies liver fibrogenesis in mice and humans. Nat Genet 2005;37:835–843. [PubMed: 15995705]

[42]. Muhlbauer M, Bosserhoff AK, Hartmann A, Thasler WE, Weiss TS, Herfarth H, et al. A novel MCP-1 gene polymorphism is associated with hepatic MCP-1 expression and severity of HCV-related liver disease. Gastroenterology 2003;125:1085–1093. [PubMed: 14517792]

[43]. Glas J, Torok HP, Tonenchi L, Schiemann U, Folwaczny C. The 2518 promotor polymorphism in the MCP-1 gene is not associated with liver cirrhosis in chronic hepatitis C virus infection. Gastroenterology 2004;126:1930–1931. [PubMed: 15188212]

[44]. Bonkovsky HL, Salek J. No role of the −2518 promoter polymorphism of monocyte chemotactic protein-1 in chronic hepatitis C. Gastroenterology 2005;129:1361–1362. [PubMed: 16230097]

[45]. Smith BC, Gorve J, Guzail MA, Day CP, Daly AK, Burt AD, et al. Heterozygosity for hereditary hemochromatosis is associated with more fibrosis in chronic hepatitis C. Hepatology 1998;27:1695–1699. [PubMed: 9620344]

[46]. Thorburn D, Curry G, Spooner R, Spence E, Oien K, Halls D, et al. The role of iron and haemochromatosis gene mutations in the progression of liver disease in chronic hepatitis C. Gut 2002;50:248–252. [PubMed: 11788568]

[47]. Hellier S, Frodsham AJ, Hennig BJ, Klenerman P, Knapp S, Ramaley P, et al. Association of genetic variants of the chemokine receptor CCR5 and its ligands, RANTES and MCP-2, with outcome of HCV infection. Hepatology 2003;38:1468–1476. [PubMed: 14647058]

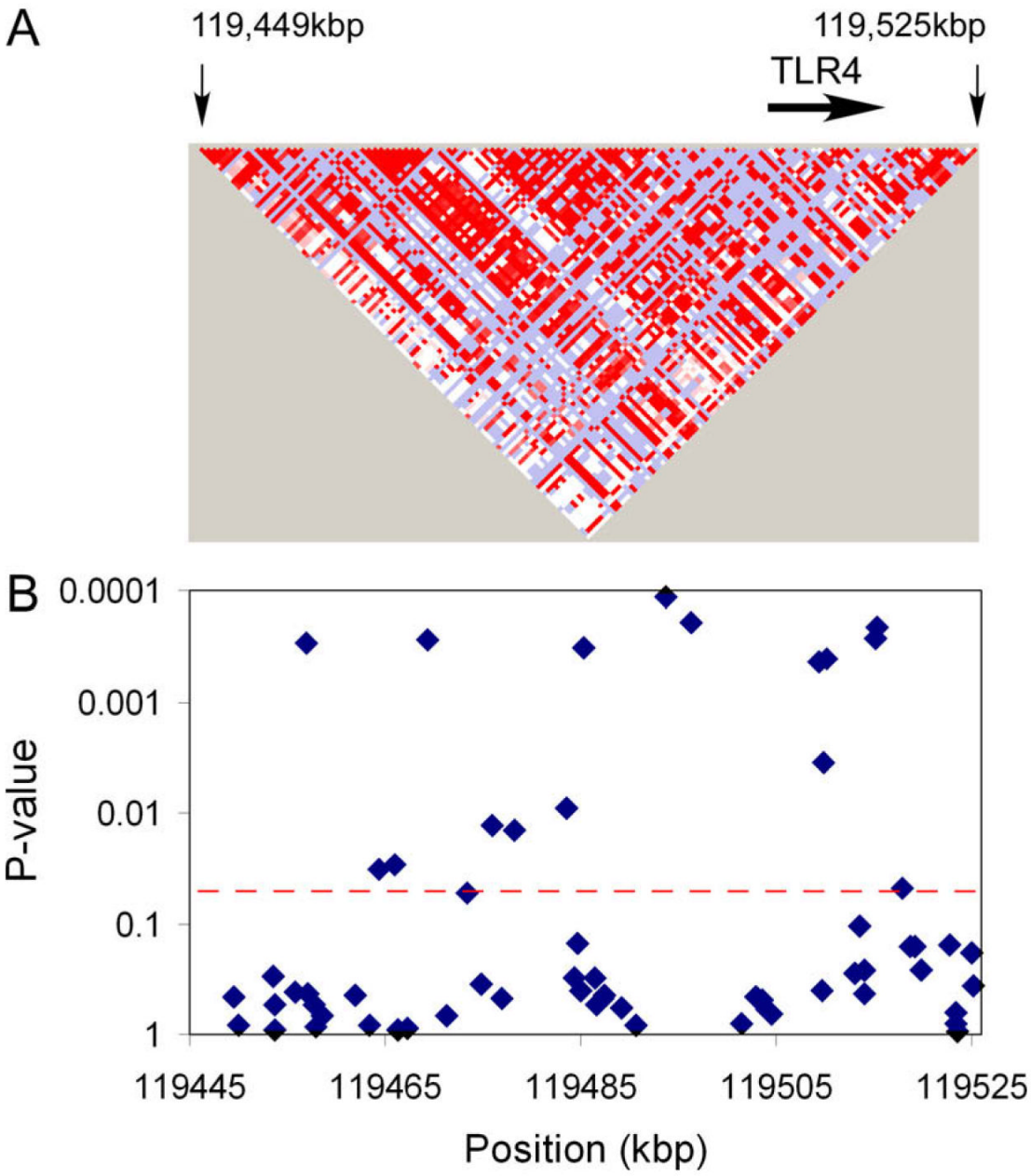

Fig. 1.

Fine-mapping analysis of the *TLR4* locus. (A) Linkage disequilibrium structure in the HapMap CEPH dataset and gene location are shown. (B) Allelic association of single SNPs with fibrosis risk. The broken line denotes $P = 0.05$.

Table 1

Fine-mapping coverage.

SNP predictor[a]	RS number	Gene	Chr	Position (bp)	Targeted region (kbp)[b]	Number of tags required[c]	Number of HapMap SNPs typed[d]	Percentage of coverage of tags	Number of non-HapMap SNPs typed	Number of total SNPs typed
SNP1	hCV25635059	*AZIN1*	8	103,910,885	159	26	55	92	16	71
SNP2	rs4986791	*TLR4*	9	119,515,423	76	34	52	88	10	62
SNP3	rs886277	*TRPM5*	11	2,396,343	97	22	22	64	22	44
SNP4	rs2290351	*AP3S2*	15	88,175,785	101	21	32	86	14	46
SNP5	rs4290029		1	222,467,263	181	29	41	69	8	49
SNP6	rs17740066	*STXBP5L*	3	122,582,973	163	27	46	74	12	58
SNP7	rs2878771	*AQP2*	12	48,638,660	20	14	9	79	14	23

[a] Based on Huang et al. [5].

[b] Region identified based on the linkage disequilibrium structure in the HapMap dataset. For the SNP6 region, marker-marker LD extends~660 kbp covering *STXBP5L* and *POLQ* genes; we only tested markers in part of the large region that includes *POLQ* gene for the purpose of determining whether *STXBP5L* or *POLQ* is more likely to be the causal gene.

[c] Determined by the HapMap tagger with $r2 > 0.8$ and minor allele frequency ≥ 0.05.

[d] Including tagging and putative functional SNPs and other SNPs such as those in high linkage disequilibrium with the original marker.

Table 2

Significant markers in high linkage disequilibrium with or independent of the *TLR4* SNP rs4986791.

Group	RS number	SNP type[a]	Case allele frequency	Control allele frequency	OR (95% CI)[b]	Allelic *P*-value	Regression *P*-value of				r^2 with rs4986791	r^2 with rs960312
							Marker adjusted for rs4986791	rs4986791 adjusted for marker	Marker adjusted for rs960312	rs960312 adjusted for marker		
1	rs4986791	Missense (T399I)	0.034	0.096	0.33 (0.18–0.61)	0.00019	N/A	N/A	0.00096	0.027	N/A	0.01
	rs4986790	Missense (D299G)	0.034	0.096	0.34 (0.19–0.62)	0.00024	0.89	0.30	0.0014	0.025	0.93	0.01
	rs10818069	Intergenic	0.044	0.115	0.35 (0.21–0.61)	0.00010	0.30	0.33	0.00056	0.025	0.74	0.01
	rs10759933	Intron	0.038	0.099	0.36 (0.2–0.64)	0.00037	0.73	0.12	0.0019	0.021	0.87	0.01
	rs7864330	Intron	0.038	0.099	0.36 (0.2–0.65)	0.00039	0.73	0.12	0.0019	0.027	0.87	0.01
	rs10818070	Intergenic	0.046	0.115	0.37(0.22–0.63)	0.00017	0.41	0.23	0.00077	0.022	0.73	0.01
	rs10448253	Intergenic	0.052	0.122	0.39 (0.23–0.66)	0.00024	0.17	0.11	0.0015	0.018	0.51	0.01
	rs12375686	Intergenic	0.051	0.121	0.39 (0.23–0.66)	0.00026	0.17	0.11	0.0015	0.026	0.51	0.01
	rs1252039	Intergenic	0.048	0.115	0.39 (0.23–0.66)	0.00029	0.63	0.13	0.0015	0.021	0.72	0.01
2	rs960312	Intergenic	0.148	0.086	1.85 (1.17–2.94)	0.0083	0.027	0.00096	N/A	N/A	0.01	N/A
	rs1329060	Intergenic	0.153	0.093	1.76 (1.12–2.76)	0.013	0.036	0.0013	0.89	0.59	0.01	0.94
	rs4837494	Intergenic	0.160	0.099	1.74 (1.12–2.7)	0.012	0.042	0.0011	0.82	0.47	0.01	0.89
	rs1927911	Intron	0.272	0.182	1.68 (1.19–2.38)	0.0029	0.016	0.0012	0.13	0.32	0.02	0.46
3	rs11536889	3'UTR	0.141	0.194	0.68 (0.47–0.99)	0.043	0.011	0.00015	0.086	0.017	0.01	0.03

[a] Missense, intron and 3′UTR SNPs are all in TLR4, and intergenic SNPs are in 5′ upstream TLR4.

[b] Sorted by effect size within each group.

Table 3

Haplotype association in the *TLR4* region with fibrosis.

Haplotype number	Haplotype	rs960312	rs4986791	rs11536889	Case frequency	Control frequency	*P*-Haplotype	*P*-Global
1	ACG	A	C	G	0.676	0.623	0.11	3.04E–05
2	ACC	A	C	C	0.141	0.194	3.15E–02	
3	ATG	A	T	G	0.034	0.096	1.63E–04	
4	GCG	G	C	G	0.148	0.086	7.20E–03	

Table 4

SNPs in the *STXBP5L* locus that are independently associated with fibrosis.

RS number	SNP type[a]	Case allele frequency	Control allele frequency	OR (95% CI)	*P*-value	Regression *P*-value of				*r*2 with rs17740066	*r*2 with rs13086038
						Marker adjusted for rs17740066	rs17740066 adjusted for marker	Marker adjusted for rs13086038	rs13086038 adjusted for marker		
rs17740066	Missense (V855I)	0.120	0.045	2.92 (1.61–5.30)	0.00026	N/A	N/A	0.0010	0.017	N/A	0.004
rs13086038[b]	Intron	0.029	0.067	0.41 (0.21–0.81)	0.0085	0.017	0.0010	N/A	N/A	0.004	N/A
rs2169302	Intergenic	0.107	0.186	0.52 (0.35–0.78)	0.0013	0.011	0.0031	0.00029	0.0018	0.016	0.007

[a] Both the intron and missense SNPs are in STXBP5L, and the intergenic SNP is in 3′ downstream of STXBP5L.

[b] In nearly perfect linkage disequilibrium with rs35827958 ($r2 = 0.97$).

Table 5

Haplotype association in the *STXBP5L* region with fibrosis.

Haplotype number	Haplotype	rs13086038	rs17740066	rs2169302	Case frequency	Control frequency	*P*-Haplotype	*P*-Global
1	TGT	T	G	T	0.746	0.702	0.18	4.49E–06
2	TGC	T	G	C	0.107	0.186	1.62E–03	
3	CGT	C	G	T	0.027	0.067	8.93E–03	
4	TAT	T	A	T	0.118	0.045	4.93E–04	