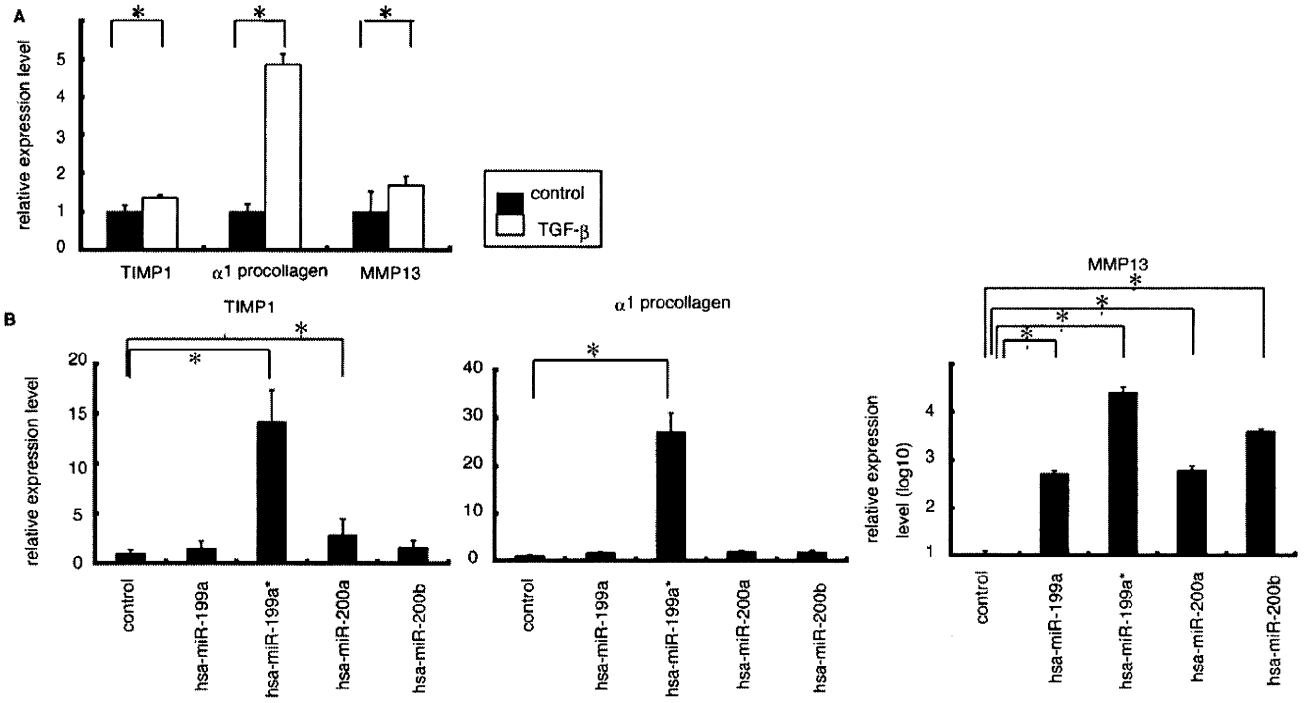

Figure 4. The relationship between expression level of miR-199 and 200 families and expression level of three fibrosis related genes. A. Administration of TGFβ in LX2 cells showed that the expression level of three fibrosis related genes were higher than that in non-treated cells. The data shown are the means+SD of three independent experiments. Asterisk was indicated to the significant difference of $p<0.05$ (two-tailed Student-t test). B. The expression levels of 3 fibrosis related genes in LX2 cells with overexpressing miR-199a, 199a*, 200a, or 200b, respectively were significantly higher than that in cells transfected with control miRNA ($p<0.05$; two-tailed Student t-test).
doi:10.1371/journal.pone.0016081.g004

Moreover miRNA expression profiling has further applications in novel anti-fibrosis therapy in CH.

Materials and Methods

Sample preparation

105 liver tissues samples from chronic hepatitis C patients (genotype 1b) were obtained by fine needle biopsy (Table S1). METAVIR fibrosis stages were F0 in 7 patients, F1 in 57, F2 in 24 and F3 in 17. Patients with autoimmune hepatitis or alcoholic liver injury were excluded. None of the patients were positive for hepatitis B virus associated antigen/ antibody or anti human immunodeficiency virus antibody. No patient received interferon therapy or immunomodulatory therapy prior to the enrollment in this study. We also obtained normal liver tissue from the Liver Transplantation Unit of Kyoto University. All of the patients or their guardians provided written informed consent, and Kyoto University Graduate School and Faculty of Medicine's Ethics Committee approved all aspects of this study in accordance with the Helsinki Declaration.

RNA preparation and miRNA microarray

Total RNA from cell lines or tissue samples was prepared using a *mir*Vana miRNA extraction Kit (Ambion, Austin, TX, USA) according to the manufacturer's instruction. miRNA microarrays were manufactured by Agilent Technologies (Santa Clara, CA, USA) and 100 ng of total RNA was labeled and hybridized using the Human microRNA Microarray Kit protocol for use with Agilent microRNA microarrays Version 1.5 and Mouse microRNA Microarray Kit protocol for use with Agilent microRNA microarrays Version 1.0. Hybridization signals were detected with a DNA microarray scanner G2505B (Agilent Technologies) and the scanned images were analyzed using Agilent feature extraction software (v9.5.3.1). Data were analyzed using GeneSpring GX 7.3.1 software (Agilent Technologies) and normalized as follows: (i) Values below 0.01 were set to 0.01. (ii) In order to compare between one-color expression profile, each measurement was divided by the 75th percentile of all measurements from the same species. The data presented in this manuscript have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE16922 (human) and accession number GSE19865 (mouse).

Real-time qPCR for human miRNA

For detection of the miRNA level by real-time qPCR, TaqMan® microRNA assay (Applied Biosystems) was used to quantify the relative expression level of miR-199a (assay ID. 002304), miR-199a* (assay ID. 000499), miR-200a (assay ID. 000502), miR-200b (assay ID. 002251), and U18 (assay ID. 001204) was used as an internal control. cDNA was synthesized using the Taqman miRNA RT Kit (Applied Biosystems). Total RNA (10 ng/ml) in 5ml of nuclease free water was added to 3 ml of 5× RT primer, 10× 1.5μl of reverse transcriptase buffer, 0.15 μl of 100 mM dNTP, 0.19 μl of RNase inhibitor, 4.16 μl of nuclease free water, and 50U of reverse transcriptase in a total volume of 15 μl. The reaction was performed for 30 min at 16°C, 30 min at 42°C, and 5 min at 85°C. All reactions were run in triplicate. Chromo 4 detector (BIO-RAD) was used to detect miRNA expression.

Animal and Chronic Mouse Liver Injury Model

Each 5 adult (8-week-old) male C57BL/6J mice were given a biweekly intra-peritoneal dose of a 10% solution of CCL_4 in olive oil (0.02 ml/g/ mouse) for the first 4 weeks and then once a week

for the next 4 weeks. At week 4, 6 or 8, the mice were sacrificed. Partial livers were fixed, embedded in paraffin, and processed for histology. Serial liver sections were stained with hematoxylin-eosin, Azan staining, Silver (Ag) staining, and Elastica van Gieson (EVG) staining, respectively. Total RNA from mice liver tissue was prepared as described previously. All animal procedures concerning the analysis of liver injury were performed in following the guidelines of the Kyoto University Animal Research Committee and were approved by the Ethical Committee of the Faculty of Medicine, Kyoto University.

Cell lines and Cell preparation

The human stellate cell lines LX-2, was provided by Scott L. Friedman. LX-2 cells, which viable in serum free media and have high transfectability, were established from human HSC lines [26]. LX-2 cells were maintained in D-MEM (Invitrogen, Carlsbad, CA, USA) with 10% fetal bovine serum, plated in 60 mm diameter dishes and cultured to 70% confluence. Huh-7 and Hela cells were also maintained in D-MEM with 10% fetal bovine serum. HuS-E/2 immortalized hepatocytes were cultured as described previously [27]. LX-2 cells were then cultured in D-MEM without serum with 0.2% BSA for 48 hours prior to TGFβ1 (Sigma-Aldrich, Suffolk, UK) treatment (2.5 ng/ml for 20 hours). Control cells were cultured in D-MEM without fetal bovine serum.

miRNA transfection

LX-2 cells were plated in 6-well plates the day before transfection and grown to 70% confluence. Cells were transfected with 50 pmol of Silencer® negative control siRNA (Ambion) or double-stranded mature miRNA (Hokkaido System Science, Sapporo, Japan) using lipofectamine RNAiMAX (Invitrogen). Cells were harvested 2 days after transfection.

Real-time qPCR

cDNA was synthesized using the Transcriptor High Fiderity cDNA synthesis Kit (Roche, Basel, Switzerland). Total RNA (2 μg) in 10.4 μl of nuclease free water was added to 1 μl of 50mM random hexamer. The denaturing reaction was performed for 10min at 65°C. The denatured RNA mixture was added to 4 μl of 5× reverse transcriptase buffer, 2 μl of 10 mM dNTP, 0.5 μl of 40U/μl RNase inhibitor, and 1.1 μl of reverse transcriptase (FastStart Universal SYBR Green Master (Roche) in a total volume of 20 μl. The reaction ran for 30 min at 50°C (cDNA synthesis), and five min at 85°C (enzyme denaturation). All reactions were run in triplicate. Chromo 4 detector (BIO-RAD, Hercules, CA, USA) was used to detect mRNA expression. The primer sequences are follows; MMP13 s; 5′-gaggctccgagaaatgcagt-3′, as; 5′-atgccatcgtgaagtctggt-3′, TIMP1 s; 5′-cttggcttctgcactgatgg-3′, as; 5′-acgctggtataaggtggtct-3′, α1-procollagen s; 5′-aacatgaccaaaaaccaaaagtg-3′, as; 5′-catt-gtttcctgtgtcttctgg-3′, and β-actin s; 5′-ccactggcatcgtgatggac-3′, as; 5′-tcattgccaatggtgatgacct-3′. Assays were performed in triplicate, and the expression levels of target genes were normalized to expression of the β-actin gene, as quantified using real-time qPCR as internal controls.

Statistical analyses

Statistical analyses were performed using Student's *t*-test; *p* values less than 0.05 were considered statistically significant. Microarray data were also statistically analyzed using Welch's test and Bonferroni correction for multiple hypotheses testing.

Supporting Information

Figure S1 Time line of the induction of chronic liver fibrosis. Upward arrow indicated administration of olive oil or CCl_4. Downward arrow indicates when mice were sacrificed.
(TIF)

Figure S2 Comparison of the expression level of miR-199 and 200 familes in several cell lines and human liver tissue. Endogenous expression level of miR-199a, 199a*, 200a, and 200b in normal liver and LX2 cell as determined by microarray analysis (Agilent Technologies). Endogenous expression level of same miRNAs in Hela, Huh-7 and, immortalized hepatocyte: HuS-E/2 by previously analyzed data [9].
(TIF)

Table S1 Clinical characteristics of patients by the grade of fibrosis.
(DOCX)

Table S2 Extracted human miRNAs related to liver fibrosis.
(DOCX)

Table S3 Corresponding human and mouse miRNAs.
(DOCX)

Table S4 Hypothetical miRNA target genes according to in silico analysis.
(DOCX)

Author Contributions

Conceived and designed the experiments: YM KS. Performed the experiments: YM HT YH NK. Analyzed the data: MT MK. Contributed reagents/materials/analysis tools: YM HT YH NK. Wrote the paper: YM MT AT FM NK TO.

References

1. Wasley A, Alter MJ (2000) Epidemiology of hepatitis C: geographic differences and temporal trends. Semin Liver Dis 20: 1–16.
2. Shepard CW, Finelli L, Alter MJ (2005) Global epidemiology of hepatitis C virus infection. Lancet Infect Dis 5: 558–567.
3. Gressner AM, Weiskirchen R (2006) Modern pathogenetic concepts of liver fibrosis suggest stellate cells and TGF-beta as major players and therapeutic targets. J Cell Mol Med 10: 76–99.
4. Nilsen TW (2007) Mechanisms of microRNA-mediated gene regulation in animal cells. Trends Genet 23: 243–249.
5. Zamore PD, Haley B (2005) Ribo-gnome: the big world of small RNAs. Science 309: 1519–1524.
6. Pillai RS (2005) MicroRNA function: multiple mechanisms for a tiny RNA? Rna 11: 1753–1761.
7. Ura S, Honda M, Yamashita T, Ueda T, Takatori H, et al. (2009) Differential microRNA expression between hepatitis B and hepatitis C leading disease progression to hepatocellular carcinoma. Hepatology 49: 1098–1112.
8. Yamamoto Y, Kosaka N, Tanaka M, Koizumi F, Kanai Y, et al. (2009) MicroRNA-500 as a potential diagnostic marker for hepatocellular carcinoma. Biomarkers 14: 529–538.
9. Murakami Y, Yasuda T, Saigo K, Urashima T, Toyoda H, et al. (2006) Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. Oncogene 25: 2537–2545.
10. Jin X, Ye YF, Chen SH, Yu CH, Liu J, et al. (2008) MicroRNA expression pattern in different stages of nonalcoholic fatty liver disease. Dig Liver Dis.
11. Ogawa T, Iizuka M, Sekiya Y, Yoshizato K, Ikeda K, et al. (2009) Suppression of type I collagen production by microRNA-29b in cultured human stellate cells. Biochem Biophys Res Commun.
12. Ji J, Zhang J, Huang G, Qian J, Wang X, et al. (2009) Over-expressed microRNA-27a and 27b influence fat accumulation and cell proliferation during rat hepatic stellate cell activation. FEBS Lett 583: 759–766.
13. Wienholds E, Kloosterman WP, Miska E, Alvarez-Saavedra E, Berezikov E, et al. (2005) MicroRNA expression in zebrafish embryonic development. Science 309: 310–311.

14. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, et al. (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. Cell 129: 1401–1414.
15. Friedman SL (2008) Hepatic fibrosis-Overview. Toxicology.
16. Roderburg C, Urban GW, Bettermann K, Vucur M, Zimmermann H, et al. (2010) Micro-RNA profiling reveals a role for miR-29 in human and murine liver fibrosis. Hepatology.
17. Venugopal SK, Jiang J, Kim TH, Li Y, Wang SS, et al. (2010) Liver fibrosis causes downregulation of miRNA-150 and miRNA-194 in hepatic stellate cells, and their overexpression causes decreased stellate cell activation. Am J Physiol Gastrointest Liver Physiol 298: G101–106.
18. Jiang J, Gusev Y, Aderca I, Mettler TA, Nagorney DM, et al. (2008) Association of MicroRNA expression in hepatocellular carcinomas with hepatitis infection, cirrhosis, and patient survival. Clin Cancer Res 14: 419–427.
19. Jiang X, Tsitsiou E, Herrick SE, Lindsay MA (2010) MicroRNAs and the regulation of fibrosis. Febs J 277: 2015–2021.
20. Marquez RT, Bandyopadhyay S, Wendlandt EB, Keck K, Hoffer BA, et al. (2010) Correlation between microRNA expression levels and clinical parameters associated with chronic hepatitis C viral infection in humans. Lab Invest.
21. Kim S, Lee UJ, Kim MN, Lee EJ, Kim JY, et al. (2008) MicroRNA miR-199a* regulates the MET proto-oncogene and the downstream extracellular signal-regulated kinase 2 (ERK2). J Biol Chem 283: 18158–18166.
22. Murakami Y, Aly HH, Tajima A, Inoue I, Shimotohno K (2009) Regulation of the hepatitis C virus genome replication by miR-199a. J Hepatol 50: 453–460.
23. Gibbons DL, Lin W, Creighton CJ, Rizvi ZH, Gregory PA, et al. (2009) Contextual extracellular cues promote tumor cell EMT and metastasis by regulating miR-200 family expression. Genes Dev 23: 2140–2151.
24. Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, et al. (2008) The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nat Cell Biol 10: 593–601.
25. Oberti F, Valsesia E, Pilette C, Rousselet MC, Bedossa P, et al. (1997) Noninvasive diagnosis of hepatic fibrosis or cirrhosis. Gastroenterology 113: 1609–1616.
26. Xu L, Hui AY, Albanis E, Arthur MJ, O'Byrne SM, et al. (2005) Human hepatic stellate cell lines, LX-1 and LX-2: new tools for analysis of hepatic fibrosis. Gut 54: 142–151.
27. Aly HH, Watashi K, Hijikata M, Kaneko H, Takada Y, et al. (2007) Serum-derived hepatitis C virus infectivity in interferon regulatory factor-7-suppressed human primary hepatocytes. J Hepatol 46: 26–36.

Genetic Epidemiology 34:543–551 (2010)

Estimation of *P*-value of MAX Test with Double Triangle Diagram for 2 × 3 SNP Case-Control Tables

Katsura Hirosawa, Takahisa Kawaguchi, Fumihiko Matsuda, and Ryo Yamada*

Center for Genomic Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan

Single nucleotide polymorphisms (SNPs) are the most popular markers in genetic epidemiology. Multiple tests have been applied to evaluate genetic effect of SNPs, such as Pearson's test with two degrees of freedom, three tests with one degree of freedom (χ^2 tests for dominant and recessive modes and Cockran-Armitage trend test for additive mode) as well as MAX3 test and MAX test, which are combination of four tests mentioned earlier. Because MAX test is a combination of Pearson's test of two degrees of freedom and two tests of one degree of freedom, the probability density function (pdf) of MAX statistics does not match pdf of χ^2 distribution of either one or two degrees of freedom. In order to calculate *P*-value of MAX test, we introduced a new diagram, Double Triangle Diagram, which was an extension of de Finetti diagram in population genetics which characterized all of the tests for 2 × 3 tables. In the diagram the contour lines of MAX statistics were consisted of elliptic curves and two tangent lines to the ellipses in the space. We normalized the ellipses into regular circles and expressed *P*-value of MAX test in an integral form. Although a part of the integral was not analytically solvable, it was calculable with arbitrary accuracy by dividing the area under pdf into finite rectangles. We confirmed that *P*-values from our method took uniform distribution from 0 to 1 in three example marginal count sets and concluded that our method was appropriate to give *P*-value of MAX test for 2 × 3 tables. *Genet. Epidemiol.* 34:543–551, 2010.

Key words: SNP; MAX test; association study; trend test

*Correspondence to: Ryo Yamada, Yoshida-konoecho, Sakyo-ku, Kyoto 606-8501, Japan. E-mail: ryamada@genome.med.kyoto-u.ac.jp
Received 9 December 2009; Revised 16 March 2010; Accepted 29 March 2010
Published online 17 August 2010 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/gepi.20510

INTRODUCTION

Genetic epidemiology has been one of the most active research fields in genetics. Since the human genome project published the reference sequence of human beings, genome-wide case-control association studies (GWAS) have been carried out on a large scale with remarkable results. In GWAS, single nucleotide polymorphisms (SNP) have been used as principal genetic markers. In individuals, SNPs have two alleles, major (M) and minor (m), and three genotypes, MM, Mm and mm. Therefore, case-control studies in GWAS consist of 2 × 3 contingency tables for the two groups (case and control) and three genotypes. The technology of molecular genetics has been progressing very rapidly and SNPs are no longer the only genetic markers to be tested in GWAS studies [Balding, 2006]. However, the importance of 2 × 3 tables has not become obsolete, because any genetic factor in DNA can be evaluated with 2 × 3 tables in case-control studies.

For 2 × 3 tables, Pearson's test of two degrees of freedom can be applied. When three categories are in order, the Cockran-Armitage trend test (CAT) of one degree of freedom is the best choice. In many cases in genetics, it is reasonable to consider that the risk of the heterozygous type (Mm) is between the risks of two homozygous types (MM and mm). Therefore, CAT has been frequently used for analyzing the additive effect, which considers the middle category as the average of the other two categories. However, dominant and recessive effects are also well known in genetics and these effects are tested frequently with 2 × 2 tables in which the risk of the heterozygous type is considered the same as the risk of two homozygous types [Balding, 2006; Cochran, 1954]. Sometimes the MAX3 test is used, which consists of three tests (CAT and dominant and recessive tests) and adopts the maximum of the three as its statistical value. In fact, MAX3 test was used with successful identification of disease-associated markers in a genome-wide association study (GWAS) [Sladek et al., 2007]. Alternatively, the MAX test or the optimal mode trend test (OMTT test) can be used [Campbell, 2005]. The MAX method or the OMTT method tests all modes between dominant and recessive including the additive mode. The MAX and OMTT methods are conceptually the same and the OMTT offered exact calculation of *P*-value of the test. In fact, all the abovementioned tests are trend tests with different types of scores [Yamada and Okada, 2009; Zheng et al., 2009]. Because the MAX test best represents the genetic hypothesis in many situations and has the highest power among these tests under the hypothesis, it seems to be the best test for 2 × 3 table tests for SNP genetic studies. However, the *P*-value of the MAX test is not analytically calculable, which is a drawback of the test. Although we previously proposed a method to calculate the exact *P*-value [Yamada and Okada, 2009], it requires a high computational load.

A method to approximation of *P*-value of the MAX test was proposed by Li et al. with good performance [Li et al., 2008]. In this paper, we introduce a diagram to display a 2×3 table test in which the contour lines of the MAX test are drawn as a combination of an ellipse and its tangent lines, and we propose a method to estimate the *P*-value of the MAX test using the diagram.

DOUBLE TRIANGLE DIAGRAM: A GEOMETRIC LAYOUT OF 2×3 TABLES IN TWO-DIMENSIONAL SPACE

DOUBLE TRIANGLE DIAGRAM AS AN EXTENSION OF THE DE FINETTI DIAGRAM

A de Finetti diagram is a ternary plot to graph the genotype frequencies of populations, where there are two alleles and the population is diploid. The diagram locates the conditions of genotype frequencies in an equilateral triangle. It is based on Viviani's theorem that at any point within the triangle, given the three lines from that point that are perpendicular to the sides of the triangle, the sum of the lengths of the lines is a fixed value, regardless of the position of the point [Cannings and Edwards, 1968].

Because the marginal counts of a 2×3 table are given, the sum of three genotypes is fixed for both cases and controls. Therefore, three genotype counts of each group can be plotted as a point in an equilateral triangle. In our double triangle diagram, two triangles for cases and controls are drawn and 2×3 tables are plotted as described below.

Let $\tau = \{\{n_{11}, n_{12}, n_{13}\}, \{n_{21}, n_{22}, n_{23}\}\}$ denote the observed table and its cell counts and $m = \{n_{1.}, n_{2.}, n_{.1}, n_{.2}, n_{.3}, n_{..}\}$ denote its marginal counts. Let $t_e = \{\{e_{11}, e_{12}, e_{13}\}, \{e_{21}, e_{22}, e_{23}\}\}$ denote the expected table and its counts under the null assumption and $\tau - t_e = \{\{d_{11}, d_{12}, d_{13}\}, \{d_{21}, d_{22}, d_{23}\}\}$ denote the difference between τ and t_e. Then, we have the following tables.

	AA	*Aa*	*aa*	*Sum*
Case	n_{11}	n_{12}	n_{13}	$n_{1.}$
Control	n_{21}	n_{22}	n_{23}	$n_{2.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.3}$	$n_{..}$

$n_{ij} \geq 0$

The difference between τ and t_e is as follows:

	AA	*Aa*	*aa*	Sum
Case	e_{11}	e_{12}	e_{13}	$n_{1.}$
Control	e_{21}	e_{22}	e_{23}	$n_{2.}$
Total	$n_{.1}$	$n_{.2}$	$n_{.3}$	$n_{..}$

$e_{ij} = n_{i.} n_{.j} / n_{..}$.

	AA	*Aa*	*aa*
Case	d_{11}	d_{12}	$d_{13} = -(d_{11} + d_{12})$
Control	$d_{21} = -d_{11}$	$d_{22} = -d_{12}$	$d_{23} = -d_{13} = d_{11} + d_{12}$

$d_{ij} = n_{ij} - e_{ij} \geq -e_{ij}$

$\sum_{i=1}^{2} d_{ij} = 0, \sum_{j=1}^{3} d_{ij} = 0.$

We introduce following coordinates:

$$(x, y) = \left(d_{12}, \frac{1}{\sqrt{3}} (d_{11} - d_{13}) \right). \quad (1)$$

These co-ordinates are based on the idea that geometric arrangement of multiple categories as below. Two categories are expressed as two vectors in opposite directions to each other. They are in one-dimensional space. Three categories are expressed as three vectors in two-dimensional space and their three tips are the vertices of a regular triangle. Four categories make a regular tetrahedron. In general, k categories make $k-1$ simplex in $k-1$ dimensional space, which has k vertices. Equation (1) is one way to give Cartesian coordinates to vertices of regular triangle (2 simplex for 3 categories). With these co-ordinates,

	AA	*Aa*	*aa*
Case	$d_{11} = -\frac{1}{2}x + \frac{\sqrt{3}}{2}y$	$d_{12} = x$	$d_{13} = -\frac{1}{2}x - \frac{\sqrt{3}}{2}y$
Control	$d_{21} = -d_{11} = \frac{1}{2}x - \frac{\sqrt{3}}{2}y$	$d_{22} = -d_{12} = -x$	$d_{23} = -d_{13} = \frac{1}{2}x + \frac{\sqrt{3}}{2}y$.

Because $n_{ij} \geq 0$ and because $d_{1j} + d_{2j} = 0$ and $d_{1j} \geq -e_{1j}$ and $d_{2j} \geq -e_{2j}$, therefore $d_{1j} \geq -e_{1j}$ and $-d_{2j} = d_{1j} \leq -e_{2j}$,

$$-e_{1j} \leq d_{1j} \leq e_{2j},$$

which can be re-written as,

$$\frac{1}{\sqrt{3}}x - \frac{2}{\sqrt{3}}e_{11} \leq y \leq \frac{1}{\sqrt{3}}x + \frac{2}{\sqrt{3}}e_{21}$$

$$-e_{12} \leq x \leq e_{22}$$

$$-\frac{1}{\sqrt{3}}x - \frac{2}{\sqrt{3}}e_{23} \leq y \leq -\frac{1}{\sqrt{3}}x + \frac{2}{\sqrt{3}}e_{13}$$

These three equations demarcate the field with three sets of parallel lines which make two equilateral triangles.

Figure 1A is the diagram of the following table, in which the case-to-control ratio is 1.5 and the total allele frequency is 0.2 and the samples are in Hardy-Weinberg equilibrium. The size of triangles is proportional to the sample size of groups and the larger and the smaller triangles represent controls and cases, respectively.

	AA	*Aa*	*aa*	*Sum*
Case	$n_{11} = 150$	$n_{12} = 520$	$n_{13} = 330$	$n_{1.} = 1{,}000$
Control	$n_{21} = 250$	$n_{22} = 680$	$n_{23} = 570$	$n_{2.} = 1{,}500$
Total	$n_{.1} = 400$	$n_{.2} = 1{,}200$	$n_{.3} = 900$	$n_{..} = 2{,}500$

DISTRIBUTION OF TEST STATISTICS IN DOUBLE-TRIANGLE DIAGRAM

CONTOUR LINES OF TEST STATISTICS

For 2×3 contingency tables in SNP case-control association studies, the above-mentioned multiple tests, Pearson's genotype test of two degrees of freedom, the three tests of one degree of freedom for additive, dominant and recessive mode, the MAX test [Zheng et al., 2009]

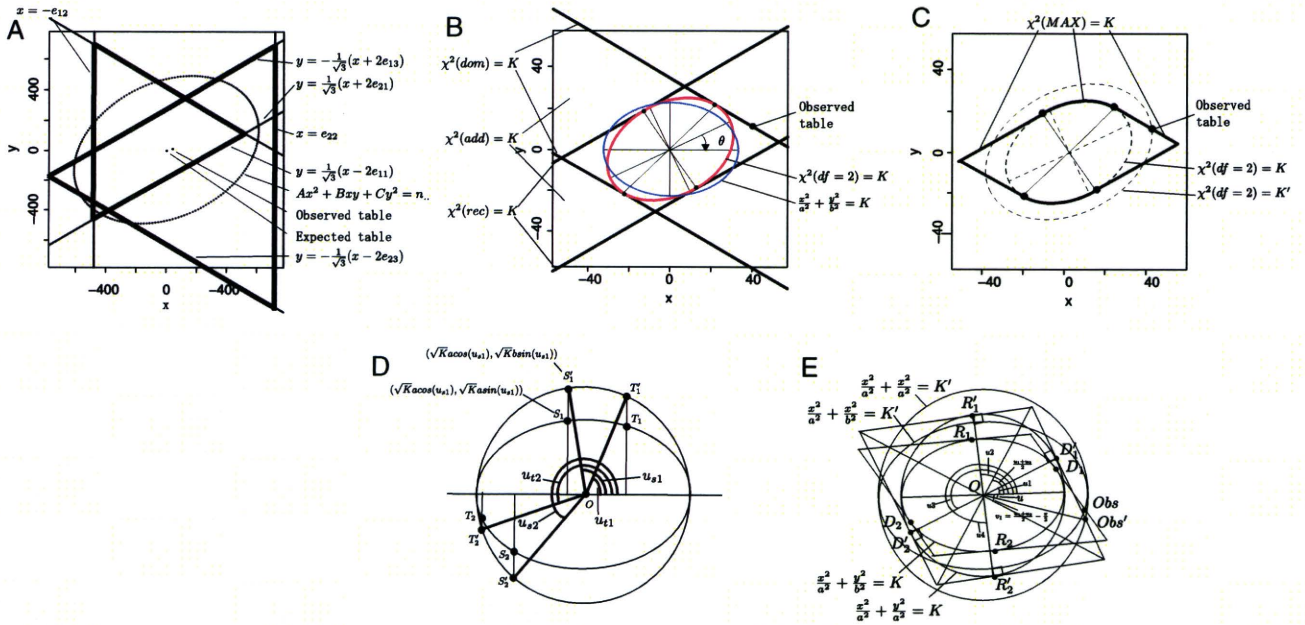

Fig. 1. Double triangle diagrams and contour lines of statistical tests. (A) The space of the tables sharing the marginals. The origin (0,0) represents the expected table and the neighboring point is the table given in the text. Six lines demarcate the space where tables exist, which share the marginals with the observed table. The overlap of two triangles is a pentagon. The dashed ellipse is the contour curve of χ^2(df2) = n. (B) The central area of (A) is enlarged. The solid ellipse of χ^2(df2) = K with its major and minor axis is shown. The solid ellipse is rotated by θ to the dashed ellipse, which is in the normalized form, $(x^2/a^2)+(y^2/b^2) = K$. Three lines, χ^2(dom) = K, χ^2(add) = K and χ^2(rec) = K are tangent to χ^2(df2) = K. The gradients of the lines are $-(1/\sqrt{3})$, 0 and $1/\sqrt{3}$, respectively. The observed table is indicated by a dot on χ^2(dom) = K and outside of the ellipse, χ^2(df = 2) = K. (C) The contour line of χ^2(MAX) = K consists of χ^2(df = K), χ^2(dom) = K and χ^2(rec) = K. In the observed table, χ^2(df2) = K' is more than K. The contour line of the ellipse, χ^2(df2) = K', is indicated by a larger dashed ellipse. Because χ^2(dom) and χ^2(df2) of the observed table are K and K', respectively, the dot of the observed table is the intersection of χ^2(dom) = K and χ^2(df2) = K'. Description on K' appears in the section "Geometric evaluation of χ^2(MAX). Therefore ignore K' when this figure was indicated earlier in the main text. (D) The ellipse $(x^2/a^2)+(y^2/b^2) = K$ is enlarged by $\frac{a}{b}$ in the y-axis direction and it becomes a circle, $(x^2/a^2)+(y^2/b^2) = K$. The coordinates of S_i, S_i', T_i and T_i' are parameterized with a, b, K, K' and u_i. The ratio of the area of the sector in the ellipse and the corresponding sector in the circle is b/a. In the figure, $u_{s1} - u_{t1} = u_{s2} - u_{t2}$. Therefore, the areas of the two sectors, S_1OT_1 and S_2T_2, are the same. (E) The ellipses, $(x^2/a^2)+(y^2/b^2) = K$ and $(x^2/a^2)+(y^2/b^2) = K'$, and the corresponding circles, enlarged by a/b in the y-axis direction, are drawn in the left and right sides, respectively.. Four tangent points on the ellipse $(x^2/a^2)+(y^2/b^2) = K$, D_1, D_2, R_1 and R_2 and their corresponding points on the circle, D_1', D_2', R_1' and R_2', are plotted. The Ds and Rs are on the lines of χ^2(dom) = K and χ^2(rec) = K, respectively. Four tangent lines to the ellipse form a parallelogram. The enlargement in the y-axis moves the lines tangent to the ellipse to the lines tangent to the corresponding circle. The tangent lines to the circle form a rhombus and the radii to D's and R's, and the tangent lines are perpendicular. Two diagonals quadrisect the rhombus. Obs and Obs′ are the point of the observed table and its corresponding point in the enlarged circle, respectively. Because the observed table's χ^2(MAX) is K and its χ^2(df = 2) is K', Obs is the intersection of the straight part of the contour line of χ^2(MAX) = K and the ellipse $(x^2/a^2)+(y^2/b^2) = K'$. Lines Obs′D_1 and OD_1 are perpendicular and angle Obs′$D_1 = u_1 - u$. The length is $D_1O = \sqrt{K}a$ and the distance is $Obs'O = \sqrt{K'}a$. Therefore, $\sqrt{K'}a/\sqrt{K}a = \cos(u_1 - u)$.

[or the OMTT test; Yamada and Okada, 2009], can be applied. As reported, these tests are all trend tests with different types of scores, as expressed below.

$$\chi^2(\text{df2}) = \max(Y^2(\{0, r, 1\}); -\infty \le r \le \infty)$$
$$\chi^2(\text{dom}) = Y^2(\{0, 0, 1\})$$
$$\chi^2(\text{rec}) = Y^2(\{0, 1, 1\})$$
$$\chi^2(\text{add}) = Y^2(\{0, 0.5, 1\})$$
$$\chi^2(\text{MAX}) = \max(Y^2(\{0, r, 1\}); 0 \le r \le 1)$$

where

$$Y^2(\{0, r, 1\}) = \frac{n_{..}^2}{n_{1.}n_{2.}} \frac{(d_{11}(-W)+d_{12}(r-W)+d_{13}(1-W))^2}{n_{.1}(-W)^2+(n_{.2}(r-W)^2+n_{.3}(1-W))^2}.$$
$$W = \frac{n_{.2} \times r + n_{.3}}{n_{..}}$$

Excluding χ^2(MAX), they are expressed with x and y as shown below.

$$\chi^2(\text{df2}) = \frac{n_{..}^2}{n_{1.}n_{2.}} \left(\frac{1}{4}\left(\frac{1}{n_{.1}}+\frac{1}{n_{.3}}\right)+\frac{1}{n_{.2}}\right) x^2 + \frac{\sqrt{3}}{2}\left(\frac{1}{n_{.3}}-\frac{1}{n_{.1}}\right)xy + \frac{3}{4}\left(\frac{1}{n_{.1}}+\frac{1}{n_{.3}}\right)y^2)$$
$$\chi^2(\text{dom}) = \frac{n_{..}^2}{n_{1.}n_{2.}} \left(\frac{1}{n_{.1}+n_{.2}}+\frac{1}{n_{.3}}\right)\left(-\frac{1}{2}x-\frac{\sqrt{3}}{2}y\right)^2 \quad (2)$$
$$\chi^2(\text{rec}) = \frac{n_{..}^2}{n_{1.}n_{2.}} \left(\frac{1}{n_{.1}}+\frac{1}{n_{.2}+n_{.3}}\right)\left(-\frac{1}{2}x+\frac{\sqrt{3}}{2}y\right)^2$$
$$\chi^2(\text{add}) = \frac{n_{..}^2}{n_{1.}n_{2.}} \frac{n_{.1}n_{.2}n_{.3}}{\frac{1}{4}\left(\frac{1}{n_{.1}}+\frac{1}{n_{.3}}\right)+\frac{1}{n_{.2}}} \left(\frac{\sqrt{3}}{2}y\right)^2$$

Let K denote χ^2(MAX) of the observed table. Figure 1B indicates the contour lines of χ^2(df2) = K, χ^2(dom) = K, χ^2(rec) = K and χ^2(add) = K. Figure 1C indicates the contour lines of χ^2(MAX) = K. The contour line of χ^2(df2) is an ellipse and the contour lines of χ^2(dom), χ^2(rec) and χ^2(add) are pairs of parallel lines. The lines of χ^2(dom) and χ^2(rec) are parallel to the lines of the triangles. The lines of χ^2(add) are horizontal. The contour lines of χ^2(MAX) consist of the elliptic curve of χ^2(df2) and the straight lines of χ^2(dom) and χ^2(rec).

The contour lines of χ^2(dom) = K, χ^2(rec) = K and χ^2(add) = K are tangent to the ellipse, which can be shown by the simple transformation of equations (not shown).

ELLIPSE NORMALIZATION

In general, the ellipse can be normalized by rotation.

The ellipse of χ^2(df2) = K is normalized to $(x^2/a^2)+(y^2/b^2) = K$, $(a \geq b)$ by rotating θ in the clockwise direction as shown below (Fig. 1B).

$$\begin{aligned}
&\frac{x^2}{a^2}+\frac{y^2}{b^2}=K\\
&a=\sqrt{\frac{2}{A+C-\sqrt{B^2+(A-C)^2}}}\\
&b=\sqrt{\frac{2}{A+C+\sqrt{B^2+(A-C)^2}}}\\
&\theta=\frac{1}{2}sin^{-1}\frac{B}{A-C}\\
&A=\frac{n_{..}^2}{n_{1.}n_{2.}}\left(\frac{1}{4}\left(\frac{1}{n_{.1}}+\frac{1}{n_{.3}}\right)+\frac{1}{n_{.2}}\right)\\
&B=\frac{n_{..}^2}{n_{1.}n_{2.}}\frac{\sqrt{3}}{2}\left(\frac{1}{n_{.3}}-\frac{1}{n_{.1}}\right)\\
&C=\frac{n_{..}^2}{n_{1.}n_{2.}}\frac{3}{4}\left(\frac{1}{n_{.1}}+\frac{1}{n_{.3}}\right)
\end{aligned} \tag{3}$$

When the ellipse is normalized, the coordinates of the points on the ellipse are given as $(\sqrt{K}a \times \cos(u), \sqrt{K}b \times \sin(u))$.

We enlarge the figure by a/b in the y-axis direction, and the ellipse becomes a regular circle. The coordinates of the points change from $(\sqrt{K}a \times \cos(u), \sqrt{K}b \times \sin(u))$ to $(\sqrt{K}a \times \cos(u), \sqrt{K}a \times \sin(u))$ (Fig. 1D).

The tangent lines to the ellipse change their gradients but remain tangent to the circle (Fig. 1E).

GEOMETRIC EVALUATION OF X^2(MAX)

The contour line of χ^2(MAX) = K consists of the tangent lines of the dominant and recessive models and the elliptic curve (see Fig. 1E). The four tangent points, R_1, R_2, D_1 and D_2 are on the ellipse and their locations are given as $(\sqrt{K}a\cos(u_i), \sqrt{K}b\sin(u_i))$. The observed table is on the larger ellipse and its location is given as $(\sqrt{K'}a\cos(u), \sqrt{K'}b\sin(u))$. The observed table is indicated as Obs at the intersection of the ellipse and the tangent line from one of the four tangent points, D_1. After enlargement in the y-axis direction, Obs and D_1 are moved to Obs′ and D_1, respectively. Because the radius and tangent line of a regular circle are perpendicular, the line Obs′$-D_1$ is perpendicular to the radius to D_1. Therefore,

$$\frac{\sqrt{K}a}{\sqrt{K'}a}=\cos(u_i-u),$$

and,

$$K'(u|u_i)=\frac{K}{(\cos(u_i-u))^2}.$$

We will use $K'(u)$ instead of $K'(u \mid u_i)$ for simplicity.

The χ^2(df2) values of points on χ^2(MAX) = K, $K'(u)$ are

$$K'(u)=\begin{cases}K & (u_1 \leq u \leq u_2, u_3 \leq u \leq u_4)\\ \min\left(\dfrac{K}{(\cos(u_i-u))^2}\right) & \text{otherwise}\end{cases} \tag{4}$$

GEOMETRIC CALCULATION OF *P*-VALUE OF MAX TEST

The area of the ellipse, $(x^2/a^2)+(y^2/b^2) = K$, is $A = \pi Kab$ (see Fig. 1D).

Assume two sectors of the ellipse, S_1OT_1 and S_2OT_2 $(S_i = (\sqrt{K}a\cos(u_{si}), \sqrt{K}b\sin(u_{si}))$ and $T_i = (\sqrt{K}a\cos(u_{ti}), \sqrt{K}b\sin(u_{ti}))$. When $u_{s1}-u_{t1} = u_{s2}-u_{t2}$, the area of the sectors, $((u_{s1}-u_{t1})/2\pi)A$ and $((u_{s2}-u_{t2})/2)A$, are equal.

Let $\Pr(x,y)$ denote the probability density function (pdf) of $\chi^2(\mathrm{df}=2)$ in two-dimensional space. The P-value of $\chi^2(\mathrm{df}=2) = K$, $P_{\mathrm{df2}}(K) = e^{-\frac{K}{2}}$ is given as,

$$P_{\mathrm{df2}}(K)=\int_{\frac{x^2}{a^2}+\frac{y^2}{b^2}\geq K}\Pr(x,y)\,dxdy. \tag{5}$$

Because $\Pr(x_i,y_i) = \Pr(x_j,y_j)$, when

$$\frac{x_i^2}{a^2}+\frac{y_i^2}{b^2}=\frac{x_j^2}{a^2}+\frac{y_j^2}{b^2}$$

the integral of $\Pr(x,y)$ in the sector S_1OT_1 is,

$$\begin{aligned}&\int_{\frac{x^2}{a^2}+\frac{y^2}{b^2}\geq K, u_{t1}\leq u\leq u_{s1}}\Pr(x,y)\frac{u_{s1}-u_{t1}}{2\pi}\,dxdy\\ &=P_{\mathrm{df2}}(K)\frac{u_{s1}-u_{t1}}{2\pi}.\end{aligned} \tag{6}$$

This can be applied to du as,

$$\int_{\frac{x^2}{a^2}+\frac{y^2}{b^2}\geq K(u_t),u_t\leq u\leq u_t+du}\Pr(x,y)\frac{du}{2\pi}dxdy=P_{\mathrm{df2}}(K(u_t))\frac{du}{2\pi}. \tag{7}$$

For the points on the contour line χ^2(MAX) = K, χ^2(df = 2) is not constant because it is a function of u, $K'(u)$ (Equation (4)). Figure 2A,B shows $K'(u)$ and $P_{\mathrm{df2}}(K(u)) = e^{-\frac{K(u)}{2}}$.

The P-value of the MAX test when χ^2(MAX) = K, $P_{\mathrm{MAX}}(K)$ is geometrically defined as,

$$P_{\mathrm{MAX}}(K)=\int_0^{2\pi}P(K'(u))\frac{du}{2\pi}. \tag{8}$$

Because the four tangent lines form a rhombus in the enlarged coordinates where the ellipse is the regular circle,

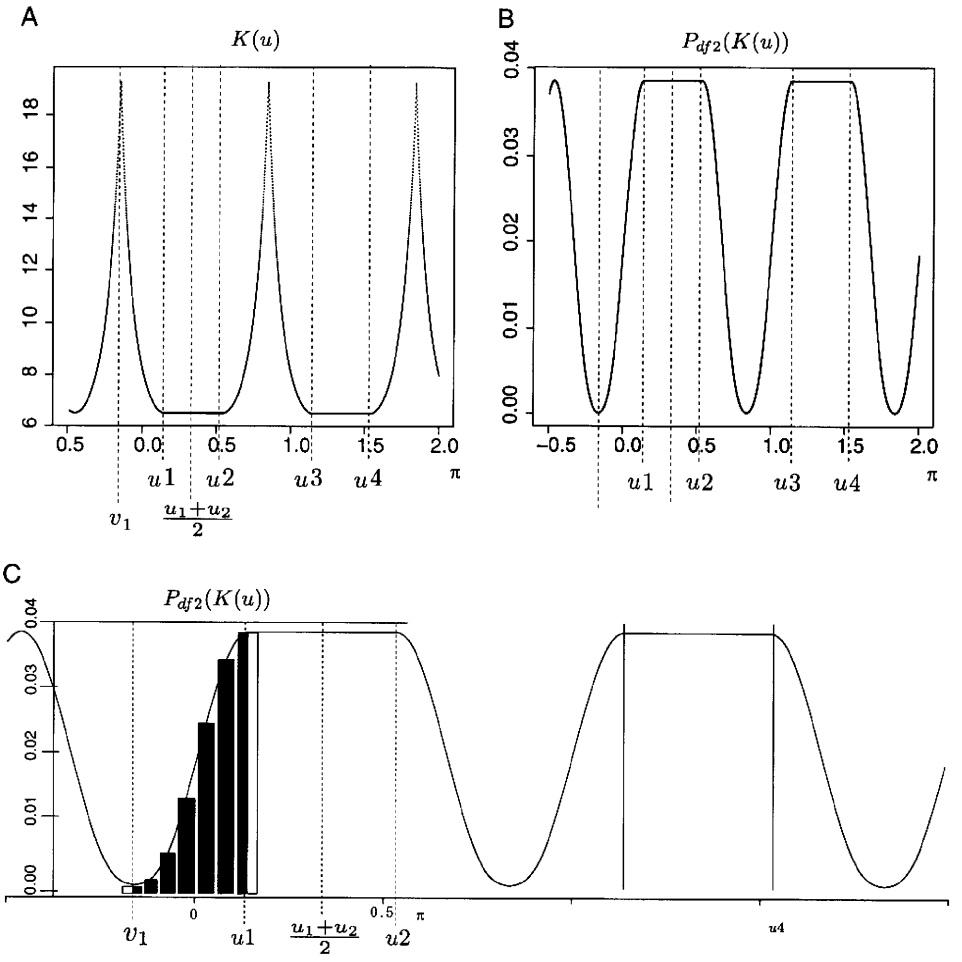

Fig. 2. Plots of χ^2(df = 2) and $P_{df2}(K)$ on the contour line of χ^2(MAX) = K to the circular angles. (A) $K'(u)$, χ^2(df2) values of points on the contour line χ^2(MAX) = K = 6.51 are plotted along the axis of $\frac{u}{\pi}$. The line is cyclic. Between u_1 and u_2 and between u_3 and u_4, the line is flat at 6.51; otherwise, $K'(u)$ is a curve larger than the value of the flat segments. (B) $P_{df2}(K(u))$ were plotted on the same horizontal axis. The flat segments correspond to the flat segments in (A). Otherwise, the line is a curve smaller than the value of the flat segments. (C) A quarter from $v_1 = \frac{u_1+u_2}{2} - \frac{\pi}{2}$ to $\frac{u_1+u_2}{2}$ of (B) is enlarged. The black bars represent the area of equation (9) for estimating $P_{MAX}(K)$. The height of the bars is $P_{MAX}(K(u)$ of the midpoint. The width of the right- and left-most black bars are half of the others.

and because a rhombus consists of four congruent regular triangles, a quarter of the rhombus can be considered. The first quarter of the rhombus of $\chi^2(\mathrm{MAX}) = K$ is parameterized with u from $v_1 = ((u_1+u_2)/2) - \pi/2$ to $v_1+(\pi/2) = (u_1+u_2)/2$ (Fig. 1E). Therefore,

$$P_{\mathrm{MAX}}(K) = 4\int_{v_1}^{v_1+\frac{\pi}{2}} P_{\mathrm{df2}}(K'(u))\frac{du}{2\pi}$$

$$= 4\int_{v_1}^{u_1} P_{\mathrm{df2}}\left(\frac{K}{(\cos(u_1-u))^2}\right)\frac{du}{2\pi} + \int_{u_1}^{v_1+4\frac{\pi}{2}} P_{\mathrm{df2}}(K)\frac{du}{2\pi}.$$

The first term of the right side is

$$4\int_{v_1}^{u_1} P_{\mathrm{df2}}\left(\frac{K}{(\cos(u_1-u))^2}\right)\frac{du}{2\pi} = \frac{2}{\pi}\int_{v_1}^{u_1} e^{-\frac{K}{2(\cos(u_1-u))^2}}du.$$

This integral cannot be analytically solved. The second term of the right side is

$$4\int_{u_1}^{v_1+\frac{\pi}{2}} P_{\mathrm{df2}}(K)\frac{du}{2\pi} = \frac{2}{\pi}P(K)\left(v_1+\frac{\pi}{2}-u_1\right) = \frac{1}{\pi}P(K)(u_2-u_1).$$

Therefore,

$$P(\chi^2(\mathrm{MAX}) = K) = \frac{1}{\pi}\left(2\int_{v_1}^{u_1} e^{-\frac{K}{2(\cos(u_1-u))^2}}\frac{du}{2\pi} + e^{-\frac{K}{2}}(u_2-u_1)\right).$$

Although the first term cannot be analytically solved, it can be calculated by summing the thin rectangles, as expressed below (Fig. 2C):

$$\frac{u_1-t_1}{2\pi N}\left(P_{\mathrm{df2}}(K'(u_1-t_1)) + P_{\mathrm{df2}}(K'(u_1)) \times \frac{1}{2} + \sum_{i=1}^{N-1} P_{\mathrm{df2}}\left(K'\left(u_1-\left(1-\frac{i}{N}\right)t\right)\right)\right). \quad (9)$$

The precision can be adjusted by changing N. In the following calculation, the first N started with 2 and it was repeatedly doubled until the difference of the estimated $P_{MAX}(K)$ by each update of N was less than $P_{df2}(K) \times 10^{-3}$. For example, when $P_{df2}(K) \times 10^{-4}$, calculation is continued until the difference between iterations χbecomes less than $10^{-4-3} = 10^{-7}$.

VALIDATION OF THE METHOD TO CALCULATE $P(\chi^2(MAX) = K)$

Previous reports have confirmed that χ^2(MAX) ranks the observed tables appropriately under the condition that r is in the hypothesized range [Yamada and Okada, 2009; Zheng et al., 2009]. Because P-values need to be observed in uniform distribution from 0 to 1 when tests are repeated under the null hypothesis, the appropriateness of the method to estimate $P_{MAX}(K)$ proposed above is confirmed by observing that the distribution of $P(\chi^2(MAX))$ is uniform for multiple tables sampled under the null hypothesis. As Equation (4) indicates, χ^2(df2) is a function of u_i, which determines the location of the tangent points. Therefore, the eccentricity of the ellipse, $\sqrt{1-(b/a)^2}$, and the angle θ affect the estimation of $P(\chi^2(MAX)) = K$. Both b/a and θ are parameterized only by n_j based on Equation (3) (details are not shown). So, we selected three marginal count sets as examples $(n_{\cdot 1}, n_{\cdot 2}, n_{\cdot 3}) = (3333, 3333, 3334), (100, 9000, 900)$ and $(9000, 100, 900)$, with $(n_{1\cdot}, n_{2\cdot}) = (5000, 5000)$, for evaluation of the method to estimate $P_{MAX}(K)$. Figure 3A shows the double triangle diagrams of the three examples. First, (3333,3333,3334) has similar values for the three categories and its diagram gives an ellipse that is almost a regular circle and the fraction where $\chi^2(MAX) = \chi(df = 2)$ is approximately one-third. In the second example, (100,9000,900) has a very large value of $n_{\cdot 2}$ and its diagram gives an ellipse that is long in the vertical axis and the fraction where $\chi^2(MAX) = \chi(df = 2)$ is very small. In the third example, (9000,100,900) has a very large value of $n_{\cdot 1}$ instead and its diagram gives an ellipse that is long in the horizontal axis and the fraction where $\chi^2(MAX) = \chi(df = 2)$ is very large. A total of 1,000 tables that had the marginal counts were randomly sampled for each example and P-P plots of $P_{MAX}(K)$ and $P_{df2}(K)$ were drawn on a linear scale and logarithmic scale (Fig. 3A). The P-P plots for P_{MAX} and P_{df2} indicate that they are uniformly distributed. There is no difference among the examples, which indicates that the difference of eccentricity or rotation of the ellipses does not affect the distributions. Figure 3B displays the relationship among P_{df2}, P_{MAX}, P_{dom}, P_{rec} and P_{add} for the 1,000 table samples of (3333,3333,3334). The correlation between P_{df2} and P_{MAX} is the strongest. Although three tests of one degree of freedom show considerable correlation with P_{MAX}, a fraction of the tables show a substantial difference.

COMPARISON OF THE POWER OF NEW METHOD WITH PEARSON'S TEST

We compared power of our method with Pearson's test and asymptotic estimation of P-value of MAX by Zang

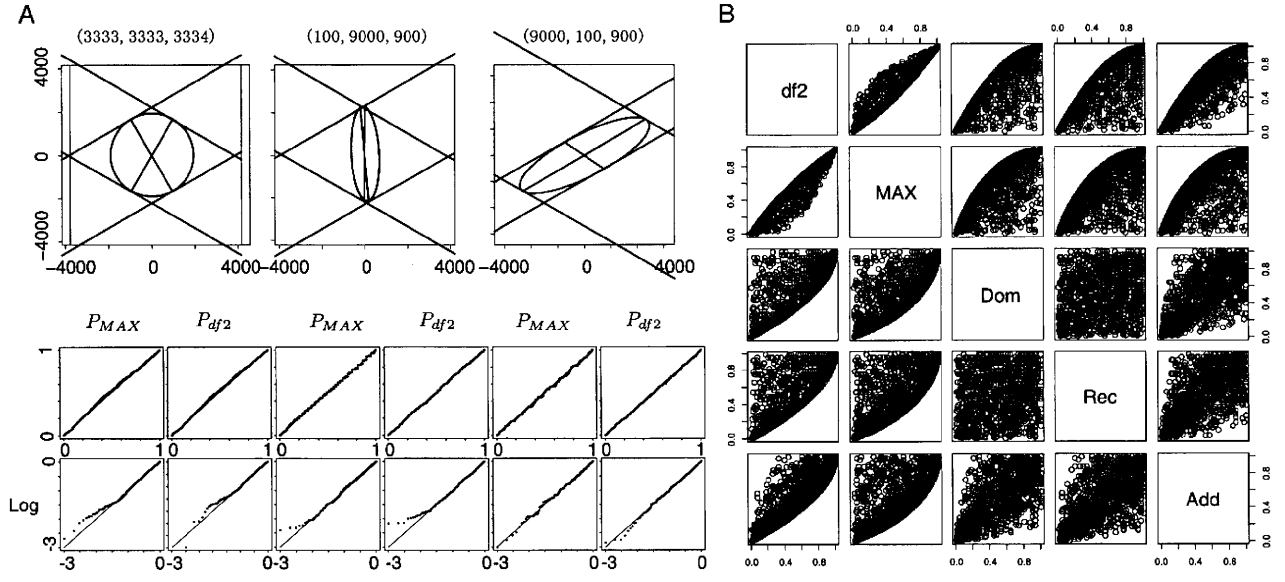

Fig. 3. Distribution of P_{MAX} for three sets of marginal counts. (A) A double triangle diagram and four P-P plots below it are drawn for three marginal counts, (3333,3333,3334), (100,9000,900) and (9000,100,900), as indicated. The two darker lines in each double triangle diagram indicate the lines connecting tangent points. They separate the sections where the contour line of $\chi^2(MAX) = K$ corresponds to the ellipse, the elliptic sections (the upper and the lower sections) and the sections where the contour line of $\chi^2(MAX) = K$ corresponds to the tangent lines (the right and left sections). The angles where the contour line of $\chi^2(MAX) = K$ is a part of ellipse is middle, narrow and wide for three examples, respectively. For each marginal count set, 1000 tables were randomly simulated under the null hypothesis. (B)Four P-P plots of P_{MAX} and P_{df2} were drawn below each diagram. The left-side plots are P_{MAX} and the right-side plots are P_{df2}. The upper plots are on a linear scale and the lower plots are on a logarithmic scale.

et al. in "Rassoc" package in CRAN (http://cran.r-project.org/), for six genetic models [Zang et al., 2010]. For each model, Hardy-Weinberg equilibrium was assumed in a population and allele frequency of risk allele in the population and prevalence of phenotype were set at 0.3 and 0.01, respectively. Genotypic relative risks for each model were given as (1.5,1.5,1), (1.5,1.25,1), (1.5,1,1) for dominant, additive and recessive models, respectively. A model between dominant and additive (half dominant model) and a model between additive and recessive (half recessive model) were also defined with (1.5,1.375,1) and (1.5,1.125,1). Relative risk of the sixth model (heterozygote-specific model) was given as (1,1.5,1). Five hundreds of cases and five hundreds of controls were randomly sampled from the population and 1,000 2×3 tables were created for each model and tested with three tests. The result was shown in Table I. Powers of proposed method and MAX3 with asymptotic *P*-value were almost identical for all models. Power of Pearson's test was less powerful than the other two for all models.

APPLICATION TO REAL GENOTYPE DATA

In response to Decision Letter, we applied our proposing method to two types of real GWAS study data. The first data were 17 SNPs that were reported with statistical significance in three papers and that were used by Li et al. to evaluate their method to approximate *P*-value of MAX test [Li et al., 2008]. The second data were 10,000 SNPs among WTCCC study for rheumatoid arthritis [The Wellcome Trust Case Control Consortium, 2007]. The ten thousands SNPs were selected from the top of the list of markers in the order of chromosomal location. For the first 17 SNPs, we applied our new method (P_{MAX}) and asymptotic *P*-value estimation of MAX3 test [($P_{MAX3asy}$ Zang et al., 2010 ("Rassoc" package in CRAN (http://cran.r-project.org/)], and the exact *P* of the MAX or the OMTT [Yamada and Okada, 2009] (P_{OMITex}) and they were shown in Table II with *P*-value based on Rhombus formula $P_{rhombus}$. In the report by Li et al., they compared $P_{rhombus}$ with empirical *P*-values of bootstraps and permutations. In this report, we adopted P_{OMITex}, instead. As shown in Table II, all methods gave similar values for all SNPs.

For 10,000 SNPs in WTCCC study, we compared P_{MAX}, $P_{MAX3asy}$ and P_{OMITex}. Figure 4A, B shows their co-plots in regular and logarithmic scales, respectively. P_{MAX} showed stronger correlation with P_{OMITex} than $P_{MAX3asy}$. Figure 4C–F shows relation of difference between P_{MAX} and P_{OMITex} or $P_{MAX3asy}$ and P_{OMITex} with allele frequency or the minimum value of 2×3 table cells. The differences between P_{MAX} and P_{OMITex} or $P_{MAX3asy}$ and P_{OMITex} were both larger when allele frequency of minor allele was smaller and the minimum value of table was smaller. When the minimum value of table was very small, any asymptotic method deviates from the exact method and this was why some of P_{MAX} were relatively more deviated from P_{OMITex}. However, excluding these exceptions, P_{MAX} tended to give closer value to P_{OMITex} than $P_{MAX3asy}$ than P_{MAXasy} regardless of the minimum of table cells. It

TABLE I. Power comparison of proposed method and Pearson's test and MAX3

	Recessive			Half recessive			Additive			Half dominant			Dominat			Heterozygout-specric		
Cut off	Proposed	Pearson's	MAX3	Proposed	Pearson's	MAX3	Proposed	Pearson's	MAX3	Proposed	Pearson's	MAX3	Proposed	Pearson's	MAX3	Proposed	Pearson's	MAX3
0.01	198	157	204	224	186	217	332	288	330	460	406	468	651	599	660	445	392	461
0.001	54	37	54	63	48	55	130	102	125	200	170	203	342	322	378	196	166	211
0.0001	14	11	15	14	12	14	41	31	45	72	60	74	189	164	191	65	55	78
0.00001	5	4	4	4	1	4	11	9	11	26	21	30	82	60	93	21	17	22
0.000001	1	0	1	0	0	0	3	3	3	8	6	9	22	18	25	8	6	8
0.0000001	0	0	0	0	0	0	1	0	1	0	0	2	9	8	11	2	2	2

Number of tables with *P* less than cut-off values in 1,000 simulations.

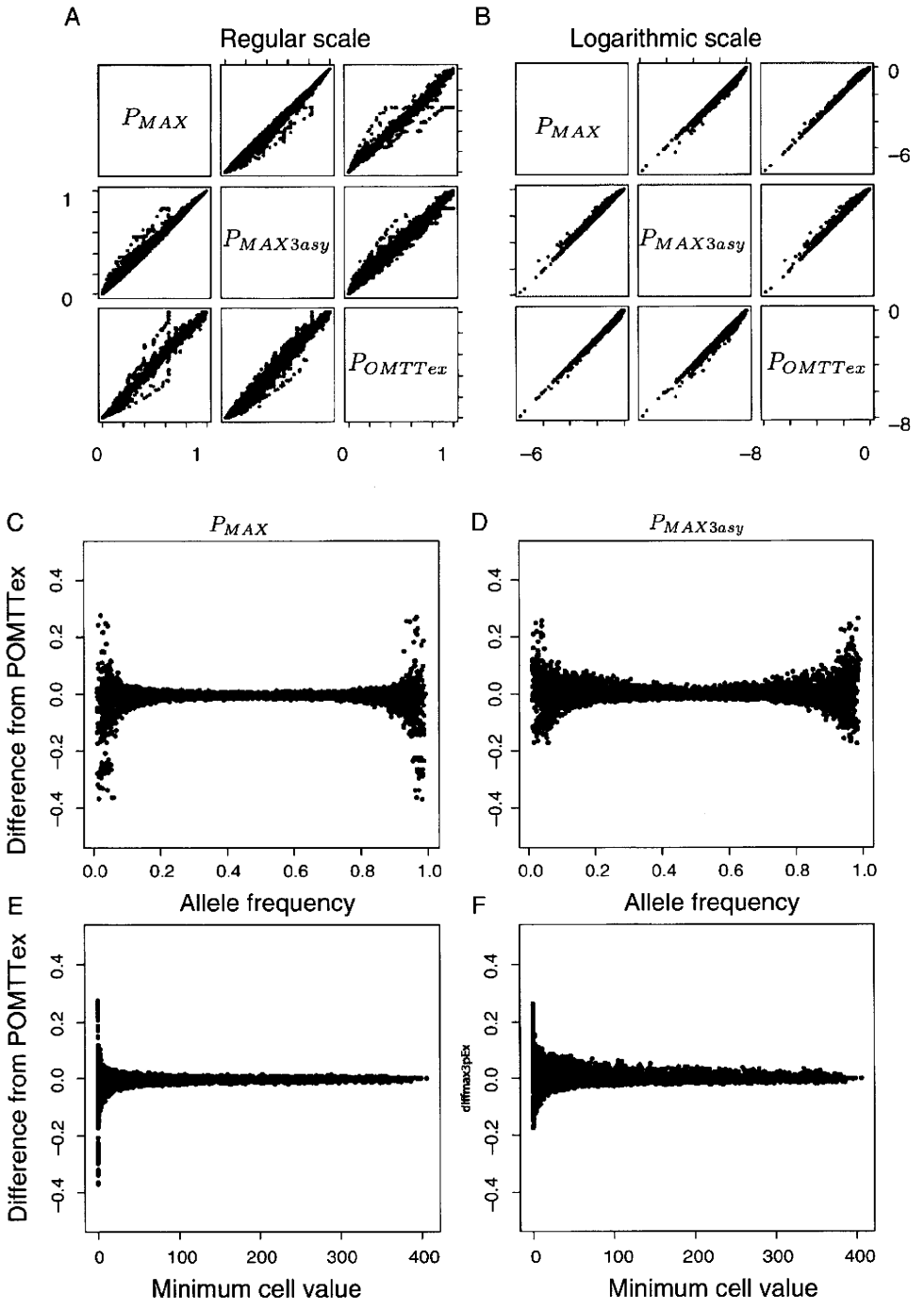

Fig. 4. P_{MAX}, $P_{MAX3asy}$ and P_{OMITex} of 10,000 SNP tables in a GWAS were compared. (A) and (B) are co-plots among P_{MAX}, $P_{MAX3asy}$ and P_{OMITex} in regular and logarithmic scale. All three were well correlated and P_{MAX} and P_{OMITex} were better. (C) and (D) Differences of P_{MAX} and $P_{MAX3asy}$ from P_{OMITex} were plotted along allele frequency. (E) and (F) Differences of P_{MAX} and $P_{MAX3asy}$ from P_{OMITex} were plotted along the minimum value of 2×3 table cells.

seemed reasonable that P_{MAX} performed relatively better with tables of low minor allele frequency or of small minimal value of cells, because the ellipses of those tables tend to have high eccentricity, to which our method was designed to correct.

DISCUSSION

In this paper, the de Finetti diagram, a diagram for genotype frequency of diallelic markers in population genetics, was applied to 2×3 contingency tables for case-control

TABLE II. *P*-values of identified SNPs in GWASs of diabetes, breast, and prostate cancers

SNPID	A11	A12	A22	B11	B12	B22	PMAX	PMAXasy	PRhombus	POMTTex
8 confirmed SNPs associated with Type 2 diabetes										
rs7903146	197	348	149	335	254	65	6.26E-19	0.00E+00	1.58E-18	3.73E-19
rs13266634	54	229	411	53	293	307	2.40E-05	1.89E-05	1.84E-05	2.27E-05
rs1111875	77	302	315	119	308	227	8.44E-06	7.04E-06	6.78E-06	7.74E-06
rs7923837	66	300	328	116	296	242	1.52E-06	2.36E-06	2.28E-06	1.43E-06
rs7480010	301	327	66	363	246	45	2.56E-05	2.24E-05	2.18E-05	2.46E-05
rs3740878	25	273	386	65	249	353	2.43E-05	1.89E-05	1.84E-05	1.94E-05
rs11037909	25	274	387	65	251	353	2.44E-05	1.89E-05	1.85E-05	2.00E-05
rs1113132	25	271	390	63	251	355	5.30E-05	4.22E-05	4.12E-05	4.11E-05
6 reported SNPs associated with breast cancer										
rs10510126	955	180	10	854	272	14	2.21E-06	1.41E-06	1.42E-06	1.63E-06
rs12505080	608	477	50	628	408	99	1.06E-04	8.46E-05	8.27E-05	1.02E-04
rs17157903	777	316	18	862	220	26	8.41E-05	6.17E-05	6.20E-05	7.56E-05
rs1219648	352	543	250	433	538	170	4.08E-06	4.99E-06	4.80E-06	3.80E-06
rs7696175	353	605	187	396	496	249	2.30E-03	2.07E-03	1.98E-03	2.37E-03
rs2420946	357	546	242	440	537	165	4.83E-06	5.34E-06	5.14E-06	4.73E-06
3 reported SNPs associated with prostate cancer										
rs1447295	25	283	864	10	218	929	1.17E-04	1.09E-04	1.10E-04	1.00E-04
rs6983267	351	598	223	277	579	301	1.95E-05	2.16E-05	2.06E-05	1.93E-05
rs7837688	861	283	27	939	206	11	9.10E-06	6.66E-06	6.67E-06	9.06E-06

association tests with SNPs, and a novel diagram, called the double triangle diagram, was proposed. Using the new diagram, test statistics of 2×3 tables were geometrically described. Given that the tests for 2×3 tables, Pearson's test, dominant test, recessive test, additive test and MAX test are all trend tests with different scores, the results were interpreted within this context. The occurrence probability distribution in the diagram space was elliptic, and the eccentricity and rotation were functions of the marginal counts. Once the distribution of the probability density was expressed in algebraic geometry, it was easy to transform the ellipse into a regular circle. In the normalized figure, integration of the probability was easy even when the integral could not be solved analytically. Subsequently, the integration of probability of the MAX test was implemented and validated by a simulation in which the geometrically estimated *P*-values of the MAX test were in a uniform distribution. Although our method performed well in terms of type I error and power and gave close value to the exact MAX test, there were limitations. In recent GWASs, tests of SNP data have to consider covariates in many cases and actually the rhombus method proposed by Li et al. [2008] was designed to be able to adjust for the covariates. However, our approach was not applicable to such conditions. Besides the handling of covariates, we evaluated our geometric approach only for 2×3 tables and no further extensions to other statistical tests of higher dimensions and more complexed data structure were investigated.

The web tool to estimate the *P*-values of the MAX test is available from the author's web site along with the R source code (http://www.genome.med.kyoto-u.ac.jp/wiki_tokyo/index.php/EllipseMAXP).

REFERENCES

Balding DJ. 2006. A tutorial on statistical methods for population association studies. Nat Rev Genet 7:781–791.

Campbell M. 2005. $\chi 2$ test for linear trend—what's that? Midwifery 32:127–130.

Cannings C, Edwards AWF. 1968. Natural selection and the de Finetti diagram. Ann Hum Gen 31:421–428.

Cochran WG. 1954. Some methods for strengthening the common $\chi 2$ tests. Biometrics 10:417–451.

Li Q, Zheng G, Li Z, Yu K. 2008. Efficient approximation of P-value of the maximum of correlated tests, with applications to genome-wide association studies. Ann Hum Genet 72:397–406.

Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P. 2007. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445:881–885.

The Wellcome Trust Case Control Consortium. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447:661–678.

Yamada R, Okada Y. 2009. An optimal dose-effect mode trend test for SNP genotype tables. Genet Epidemiol 33:114–127.

Zang Y, Wing F K, Zheng G. 2010. Simple algorithms to calculate asymptotic null distributions of robust tests in case-control genetic association studies in R. J Stat 33:1–24.

Zheng G, Joo J, Yand Y. 2009. Pearson's test, trend test, and MAX are all trend tests with different types of scores. Ann Hum Genet 73:133–140.

Interferon γ receptor 2 gene variants are associated with liver fibrosis in patients with chronic hepatitis C infection

Bertrand Nalpas,[1] Roubila Lavialle-Meziani,[2] Sabine Plancoulaine,[3,4] Emmanuelle Jouanguy,[3,5] Antoine Nalpas,[1] Mona Munteanu,[6] Frederic Charlotte,[6] Brigitte Ranque,[3,4] Etienne Patin,[3,4] Simon Heath,[2] Hélène Fontaine,[1,4] Anaïs Vallet-Pichard,[1,4] Dominique Pontoire,[7] Marc Bourlière,[8] Jean-Laurent Casanova,[3,4,9,10] Mark Lathrop,[2] Christian Bréchot,[11] Thierry Poynard,[6] Fumihiko Matsuda,[2,12] Stanislas Pol,[1,4] Laurent Abel[3,4,9]

► Supplementary tables are published online only. To view these files please visit the journal online (http://gut.bmj.com).

For numbered affiliations see end of article.

Correspondence to
Dr Laurent Abel, Laboratoire de Génétique Humaine des Maladies Infectieuses, Université Paris Descartes-INSERM U550, Faculté de Médecine Necker, 156 rue de Vaugirard, 75015 Paris, France;
laurent.abel@inserm.fr

Accepted 19 March 2010
Published Online First 29 June 2010

ABSTRACT

Background Only a minority of patients with chronic hepatitis C virus (HCV) infection develops severe liver fibrosis, a process that may be controlled by human genetic factors.

Objective To investigate the role of 384 single nucleotide polymorphisms (SNPs) located in 36 candidate genes related to the fibrogenesis/fibrolysis process.

Methods Patients with chronic HCV infection were gathered from two French cohorts (prospectively and retrospectively). The overall sample consisted of 393 HCV-infected subjects without known risk factors for fibrosis progression, including 134 patients with severe liver fibrosis and 259 without severe fibrosis.

Results Only two SNPs in strong linkage disequilibrium (LD) in the interferon γ receptor 2 gene (*IFNGR2*) were significantly associated with liver fibrosis in both the prospective and the retrospective samples. The strongest association ($p=8\times10^{-5}$) was observed with the G/A SNP rs9976971 with an OR of severe fibrosis for AA versus AG or GG subjects at 2.95 (95% CI 1.70 to 5.11). This effect was higher ($p=9\times10^{-7}$) when taking into account the time of follow-up, and the hazard ratio of progression towards severe fibrosis for AA patients was 2.62 (1.76 to 3.91). Refined sequencing and analysis of the *IFNGR2* region identified two additional variants in strong LD with rs9976971. No haplotypes derived from this cluster of four variants provided stronger evidence for association than rs9976971 alone.

Conclusions This identification of a cluster of four *IFNGR2* variants strongly associated with fibrosis progression in chronic HCV infection underlines the role of IFNγ in the development of liver fibrosis that may pave the way for new treatments.

Significance of this study

What is already known about this subject?

- Only a minority of patients with chronic HCV infection develops severe liver fibrosis.
- Viral and non-genetic host factors cannot account for the variability in the rate of progression towards liver fibrosis in patients with chronic HCV infection.
- There is accumulating evidence for the role of host genetic factors in the development of liver fibrosis, although these factors are as yet largely elusive.

What are the new findings?

- We investigated the role of 384 single nucleotide polymorphisms (SNPs) located in 36 candidate genes related to the fibrogenesis/fibrolysis process.
- We identified a single cluster of variants in the interferon γ receptor 2 gene (*IFNGR2*) strongly associated with progression to severe fibrosis.
- This association was replicated in an independent population sample, and was stronger when taking into account the time of follow-up from contamination to liver biopsy with a hazard ratio of developing severe fibrosis estimated as 2.62 (1.76 to 3.91) for the subjects homozygous for the predisposing allele.

How might it impact on clinical practice in the foreseeable future?

- As IFNγ is an anti-fibrogenic cytokine available for clinical purposes, our study may open new therapeutic avenues for the prevention of cirrhosis in HCV-infected patients.

INTRODUCTION

Hepatitis C virus (HCV) infection is a major public health concern world wide with an estimated 170 million people infected.[1 2] The natural history of patients with HCV chronic infection is characterised by a highly variable disease progression.[1–5] Most subjects never develop cirrhosis during their lifetime while the remaining patients are considered 'rapid fibrosers' and may develop severe fibrosis in <20 years.[1 5–7] While viral factors such as HCV genotypes or viral load do not seem to influence progression, several host factors such as gender (male), age at infection (>40 years old), alcohol consumption (>50 g/day), obesity and its related metabolic disorders, co-infections (in particular by HIV) are associated with the development of fibrosis.[4 5 8 9] However, these factors can account for only a minority of the variability in the rate of

progression, and a number of studies have investigated the role of genetic host factors.[5 10] Most of these studies have tested a single or a few candidate genes, and have not produced conclusive results as they were not clearly replicated.[10 11] A recent study investigated a large number (>24 000) of putative functional single nucleotide polymorphisms (SNPs) and identified two variants associated with severe fibrosis.[12] In a subsequent study which used the same SNPs and focused on Caucasian patients with well-characterised liver histology, a different panel of seven SNPs was found to predict the risk of developing cirrhosis, and this panel needs to be validated in prospective studies.[13]

Liver fibrosis is the consequence of a generalised wound-healing response of hepatic tissue against repeated injury, which results in the formation of scar tissue instead of normal parenchyma.[14] This process is characterised by an imbalance between matrix synthesis (ie, fibrogenesis) and matrix degradation (ie, fibrolysis), and leads to an accumulation of a large variety of matrix proteins, including collagens, proteoglycans and glycoproteins.[15 16] Activated hepatic stellate cells, portal fibroblasts and myofibroblasts of bone marrow origin are the major collagen-producing cells in the injured liver.[14] A number of molecules and regulatory pathways are involved in this complex process of fibrogenesis/fibrolysis.[14–16] Therefore, we hypothesised that variations in human genes involved in fibrogenesis/fibrolysis might account for interindividual variability in development of liver fibrosis among patients with chronic HCV infection. In this work we focused on the role of polymorphisms located in a panel of 36 genes (table 1) encoding either enzymes involved in extracellular matrix turnover (matrix metalloproteinases and their inhibitors) or some cytokines known to exhibit profibrogenic (transforming growth factor β (TGFβ) and related molecules) or anti-fibrogenic (interferon γ (IFNγ) and its receptors) activity. In addition to classic case–control analysis, in this study we analysed the influence of SNPs directly on the time of progression by restricting our sample to patients with a known presumed date of infection and without known risk factors of fibrosis progression such as chronic alcohol intake, associated infections or metabolic syndrome.

Table 1 List of the genes investigated in the association study

Gene name	Abbreviation
Alpha-2-macroglobulin	A2M
Angiotensinogen	AGT
Interferon gamma	IFNG
Interferon gamma receptor 1	IFNGR1
Interferon gamma receptor 2	IFNGR2
Keratin 8	KRT8
Latent transforming growth factor beta binding protein 1	LTBP1
Latent transforming growth factor beta binding protein 2	LTBP2
Latent transforming growth factor beta binding protein 3	LTBP3
Latent transforming growth factor beta binding protein 4	LTBP4
Matrix metalloproteinase 1	MMP1
Matrix metalloproteinase 2	MMP2
Matrix metalloproteinase 3	MMP3
Matrix metalloproteinase 7	MMP7
Matrix metalloproteinase 8	MMP8
Matrix metalloproteinase 9	MMP9
Matrix metalloproteinase 10	MMP10
Matrix metalloproteinase 11	MMP11
Matrix metalloproteinase 12	MMP12
Matrix metalloproteinase 13	MMP13
Matrix metalloproteinase 14	MMP14
Matrix metalloproteinase 15	MMP15
Matrix metalloproteinase 16	MMP16
Matrix metalloproteinase 17	MMP17
Matrix metalloproteinase 24	MMP24
Matrix metalloproteinase 25	MMP25
Transforming growth factor beta 1	TGFB1
Transforming growth factor beta 2	TGFB2
Transforming growth factor beta 3	TGFB3
Transforming growth factor beta receptor I	TGFBR1
Transforming growth factor beta receptor II	TGFBR2
Transforming growth factor beta receptor III	TGFBR3
Tissue inhibitor of metalloproteinase 1	TIMP1
Tissue inhibitor of metalloproteinase 2	TIMP2
Tissue inhibitor of metalloproteinase 3	TIMP3
Tissue inhibitor of metalloproteinase 4	TIMP4

PATIENTS AND METHODS

Patients

We recruited adult Caucasian patients (>18 years of age) with chronic HCV infection defined as the presence of circulating HCV RNA tested by reverse transcriptase PCR. The patients were gathered in two steps. First, we conducted a prospective enrolment of patients from the hepatology units of Necker Hospital in Paris and St Joseph Hospital in Marseille (sample A). The criteria for the inclusion of patients were (*a*) an available liver biopsy before any treatment; (*b*) a known presumed date of HCV acquisition (date of the first exposure to blood products, or of beginning of intravenous drug (IVD) use); (*c*) a low alcohol consumption (less than three or less than two standard drinks a day for men or women, respectively); (*d*) absence of co-infection with HIV or hepatitis B virus; (*e*) absence of any coexisting chronic liver disease or hepatocellular carcinoma. Clinical risk factors, history of HCV acquisition and of alcohol consumption (assessed using time-line follow-back interview) were recorded through face-to-face interviews conducted by doctors trained in addiction problems. In a second step, we gathered additional patients from an existing cohort from the hepatology unit of Pitié-Salpêtrière Hospital in Paris (sample B). The inclusion criteria were the same as for sample A except that the presumed date of infection was not known for all patients. The study was approved by the appropriate institutional review boards, and written informed consent was obtained from all patients.

Most of the enrolled patients had been followed up for a number of years in the corresponding clinics and had had a liver biopsy at the time of their first evaluation. The stage of fibrosis was assessed from liver biopsy samples using METAVIR units, and graded on a five-point scale from 0 to 4.[17] For this study, in order to optimise the phenotype definition, we excluded patients with grade 2, and retained only patients with grades 0 or 1 (F0–1 patients) referred to as having no fibrosis, and patients with grades 3 or 4 (F3–4) referred to as having severe fibrosis. For patients who had had several biopsies, we used either the first biopsy specimen showing severe fibrosis (F3–4 patients) or the last biopsy showing no fibrosis without any treatment (for F0–1 patients). The duration of infection was estimated from the presumed year of HCV acquisition (eg, first exposure to blood transfusion, or to IVD use) to the year of the relevant biopsy.

Genotyping and sequencing methods

In this study we focused on the role of polymorphisms located in a panel of genes encoding either enzymes involved in

extracellular matrix turnover (matrix metalloproteinases and their inhibitors) or cytokines that are believed to have a profibrogenic (TGFβ and related molecules) or an anti-fibrogenic (IFNγ and its receptors) activity. A total of 36 genes were selected and are shown in table 1. A selection of SNPs within each gene, totalling 384 SNPs (list in online supplementary table 1), was made using data from the first public release of HapMapII. For each gene, all HapMap SNPs in the region including the gene and the 10 kb flanking regions were initially considered. SNPs with minor allele frequency <5% or with low predicted quality for genotyping (calculated by Illumina) were filtered out, and pairwise linkage disequilibrium (LD) was estimated (from the HapMap data) between all pairs of remaining SNPs within each gene. The 384 SNP panel was then selected such that no two SNPs in the same gene had an estimated $r^2 \geq 0.8$ or were <60 bp apart.

All DNA samples were extracted from whole blood, and subjected to rigorous quality control to check for fragmentation and amplification. All SNPs were genotyped on an ultra-high throughput Illumina platform. This platform uses the GoldenGate assay followed by a bead-based technology to resolve individual SNP genotypes.[18] Discovery of SNPs within an *IFNGR2* region of ~12.3 kb from 33 689 894 to 33 702 179 bps on chromosome 21 was performed by exhaustive sequencing (figure 1). The sample consisted of 32 French Caucasian subjects (men and women with no disease history) from the Epidemiological study on the Genetics and Environment of Asthma.[19] The sample size of 32 allowed us to detect SNPs with a minor allele frequency of at least 5% with a probability of 96%. Sequencing reactions were performed with the Dye Terminator method using an ABI PRISM 3730 DNA Analyser (Applied Biosystems, Foster City, California, USA). Sequence alignment and SNP discovery were performed with Genalys software, developed by the Centre National de Génotypage (CNG).[20]

Statistical methods

Association between severe fibrosis and the panel of SNPs was first tested in sample A by a case–control analysis using the genotypic test statistic (two degrees of freedom): the cases are HCV-infected patients with severe fibrosis and the controls are infected patients without severe fibrosis. When a type I error of 0.02 was used, our initial sample A had a power of 80% for detecting a polymorphism with an additive effect, providing an odds ratio (OR) for heterozygosity of two and having a frequency >0.09. For SNPs showing association at p<0.02, we then tested association by a survival analysis approach using a Cox model: we considered as starting points the estimated ages at infection, and as end points either the first biopsy showing severe fibrosis (failure time) or the last biopsy showing absence of severe fibrosis in the absence of any treatment (censored time). For all these analyses, we determined the genetic model (dominant/additive/recessive) providing the best fit to the data. SNPs showing the most interesting results (p<0.02 in the case–control study and p<0.05 in the survival analysis) were then tested for replication in sample B.

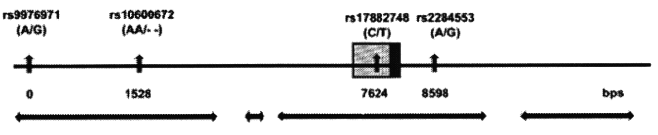

Figure 1 Schematic representation of the chromosome 21 region ranging from 33 689 800 to 33 702 200 bps, and including the 5′ region, exon 1 and part of intron 1 of *IFNGR2*. Exon 1 ranges from 33 697 072 to 33 697 792 bp with an untranslated and a translated part shown as a hatched and a solid box, respectively. Horizontal arrows indicate the regions covered by direct sequencing. Three segments could not be sequenced for technical reasons (33 694 486–33 696 000, 33 696 161–33 696 390 and 33 699 114–33 700 075 bp). The four single nucleotide polymorphisms associated with severe fibrosis are indicated by vertical arrows with distance in base pairs provided from the position of rs9976971, which is located at 33 689 967 bp.

We tested for heterogeneity of the association results according to different criteria, such as gender, mode of infection (blood transfusion/IVD use/others), viral genotypes (one and four vs others), age at infection (≤20 years vs >20 years). Under the hypothesis of homogeneity of association, twice the difference between the likelihood of the whole sample and the summed likelihoods of the subsamples (eg, the two subsamples of men and women) is asymptotically distributed as a χ^2 with one degree of freedom. All statistical analyses were performed using different procedures (FREQ, LOGISTIC, PHREG) implemented in SAS software version 8.2 (SAS Institute). Pairwise LD between SNPs was assessed by determining the r^2 coefficient using the Haploview software.[21] Haplotype analysis was conducted using the THESIAS software (http://www.genecanvas.org, accessed 5 May 2010).[22]

RESULTS

Description of the two samples

Samples A and samples B consisted of a total of 267 (103 F3–4 patients with severe fibrosis and 164 F0–1 patients without severe fibrosis), and 126 (31 F3–4 and 95 F0–1 patients) patients with chronic HCV infection, respectively. The main features of the overall sample including 393 patients are shown in table 2. There was an overall excess of women (58.5% vs 41.5%, $p=8\times10^{-4}$), which might be explained, in part, by the inclusion criterion of low alcoholic consumption, but there was no significant difference (p=0.24) in the distribution of gender according to fibrosis status. In the 364 patients with reliable HCV acquisition data, the overall distribution of modes of infection was different (p=0.0004) according to the fibrosis status. The proportion of patients infected by blood transfusion was higher in patients with severe fibrosis (57.5%) than in F0–1 patients (41.3%), while the reverse was observed for IVD users (43% in F0–1 patients vs 22% in F3–4 patients). F3–4 patients were significantly older (p=0.004) at infection (mean 29.9 years, range 0.1–73.2 years) than F0–1 patients (25.6 years, 0.1–70.8 years), and had a longer ($p=10^{-4}$) duration of HCV infection at the time of biopsy (22.8 years, 0.7–49.6 years vs 18.9 years, 0.2–49.6 years). Finally, the distribution of viral genotypes combined as usual in three main groups according to their sensitivity to anti-viral treatment[23] was not significantly different between F0–1 and F3–4 patients.

Four variants are associated with fibrosis in the first sample

Of the 384 SNPs, 16 SNPs could not be genotyped (supplementary table 1), and we excluded an additional five SNPs because either they showed deviations (p<0.005) from Hardy–Weinberg equilibrium (three SNPs) or they had a minor allele frequency <0.02 (two SNPs). The 363 remaining SNPs all showed a genotyping success >96%, and were used for association analysis. Association between severe fibrosis and the panel of 363 SNPs was first tested in sample A by a classic case–control analysis where cases were the HCV-infected patients with severe fibrosis and controls were the infected patients without severe fibrosis (supplementary table 1). A total of nine

Gut 2010;**59**:1120–1126. doi:10.1136/gut.2009.202267

Table 2 Main features of the HCV chronically infected patients in the whole sample

	METAVIR fibrosis score		
	F0–1	F3–4	Total
Number of patients	259	134	393
Sex ratio (male/female)	0.65	0.84	0.71
Mode of infection			
Blood transfusion	98	73	171
IVD	102	28	130
Others	37	26	63
Unknown	22	7	29
Age at contamination (years)*	25.6 (12.1)†	29.9 (14.0)	27.1 (12.9)
Duration of infection (years)*	18.9 (8.2)	22.8 (9.6)	20.3 (8.9)
Viral genotypes			
1A, 1B, 4	157	96	253
2, 5	27	8	35
3	49	19	68
Unknown	26	11	37

*Data for the 364 patients with known presumed dates of acquisition.
†mean (SD).
IVD, intravenous drug.

SNPs, including two in *IFNGR2* and three in *MMP16*, provided evidence for association with a p value <0.02, and were investigated further. Of these nine SNPs, four significantly influenced (p<0.05) the rate of progression towards severe fibrosis when performing a survival analysis (table 3 and supplementary table 2). The effects of the two *IFNGR2* SNPs, rs9976971 and rs2284553, which already yielded the lowest p values in the case–control study ($p=3\times10^{-4}$ and $p=8\times10^{-4}$, respectively), were even more significant ($p=2\times10^{-5}$ and $p=8\times10^{-5}$, respectively) using the survival analysis. Conversely, the effects of the two other SNPs (one in *MMP16*, and one in *TGFBR2*) were slightly lower when accounting for time of progression. For both rs9976971 and rs2284553 (which are G/A SNPs), the risk allele was the minor allele A and the best fitting genetic model was recessive—that is, subjects who were AA homozygous were predisposed to severe fibrosis as compared with AG and GG subjects. These two SNPs were in strong LD ($r^2=0.79$), and multivariate analysis confirmed that the results observed with rs9976971 and rs2284553 reflect a single signal.

Two *IFNGR2* variants show evidence for replication in the second sample

The four SNPs providing evidence for association both in case–control (at p<0.02) and survival (at p<0.05) were tested in sample B (table 3). Only the two *IFNGR2* SNPs showed evidence for replication with the same risk allele. As in sample A, the survival analysis was more powerful than the case–control approach, leading to a significant effect in sample B for rs9976971 (p=0.011). A similar trend, although not significant (p=0.13), was observed for rs2284553. When samples A and B were combined (table 4), the overall effect of rs9976971 in the case–control design was highly significant ($p=8\times10^{-5}$), and the OR (95% CI) of presenting severe fibrosis for AA subjects as compared with AG or GG subjects was 2.95 (1.70 to 5.11). This effect was much stronger ($p=9\times10^{-7}$) in the survival analysis design, and the HR (95% CI) of progressing towards severe fibrosis for AA subjects as compared with AG or GG subjects was 2.62 (1.76 to 3.91) (figure 2).

Consistent with this result taking into account the duration of infection, we observed that the effect of rs9976971 in the classical case–control design was much stronger (OR=4.46 (2.28 to 8.72)) in the 57 F3–4 patients with rapid progression (≤20 years of infection) than in the 69 F3–4 patients with slow progression (>20 years of infection, OR=2.10 (1.04 to 4.25)). Even if we considered the fact that these results were obtained in a one-step strategy (without the use of a replication sample), and that we applied the classical and stringent Bonferroni correction for multiple testing (assuming we have tested the 363 SNPs in the whole sample), the corrected p values for rs9976971 remained significant and equal to 0.029 and 0.0003 in the case–control and the survival analysis, respectively. These results indicate that one SNP in *IFNGR2*, either rs9976971 or another variant in strong LD with it, strongly influences the rate of progression towards severe fibrosis in patients chronically infected by HCV.

Search for other polymorphisms in linkage disequilibrium with the two *IFNGR2* variants

Next, we searched for other variants in strong LD with rs9976971 along three lines. First, we looked for long range LD (from 33 380 000 to 34 000 000 bp) using the European population of the HapMap database, NCBI build 36 (http://www.hapmap.org/, accessed 5 May 2010). Substantial LD (r^2 ranging from 0.3 to 0.48) with rs9976971 was observed with three clusters of SNPs. We genotyped one tag-SNP within each cluster (rs2834208, rs13047599, and rs7279549), and no association with severe fibrosis (p>0.2) was observed with any of these tag-SNPs. No other SNPs showed r^2 >0.3 within this interval, and, in particular, all SNPs between 33 524 000 and 33 655 000 bp where the genes *IFNAR1* and *IFNAR2* encoding interferon α/β receptors are located, provided r^2 <0.11 with rs9976971. Second, we explored the SeattleSNPs variation discovery resource database as it contained results for *IFNGR2*

Table 3 Results in both samples for the four single nucleotide polymorphisms (SNPs) providing the strongest evidence for association with severe fibrosis in sample A (ie, p<0.02 in case–control study and p<0.05 in the survival analysis). The full distribution of genotypes for these four SNPs is shown in online supplementary table 2

				Sample A				Sample B			
				Case/control (n=103/164)		Survival analysis (n=267)		Case/control (n=31/95)		Survival analysis (n=97)	
Marker	Gene	Risk allele (frequency)	Model*	OR (95% CI)†	p	HR (95% CI)‡	p	OR (95% CI)	p	HR (95% CI)	p
rs2284553	IFNGR2	A (0.41§)	rec	3.10 (1.57 to 6.13)	8×10^{-4}	2.41 (1.54 to 3.79)	8×10^{-5}	1.02 (0.26 to 4.05)	-¶	3.01 (0.66 to 13.7)	0.13
rs9976971	IFNGR2	A (0.44)	rec	3.22 (1.68 to 6.18)	3×10^{-4}	2.54 (1.64 to 3.93)	2×10^{-5}	2.04 (0.68 to 6.17)	0.19	3.56 (1.26 to 10.1)	0.011
rs2664357	MMP16	C (0.27)	add	1.86 (1.24 to 2.79)	0.002	1.47 (1.12 to 1.93)	0.006	0.62 (0.32 to 1.21)	0.16	0.81 (0.40 to 1.65)	–
rs9831477	TGFBR2	T (0.58)	dom	2.93 (1.35 to 6.38)	0.005	2.27 (1.14 to 4.50)	0.016	1.58 (0.49 to 5.08)	–	0.97 (0.32 to 2.91)	–

*Genetic model (recessive (rec), additive (add), or dominant (dom)) for the risk allele.
†OR (with 95% CI) under the corresponding genetic model.
‡HR (with 95% CI) under the corresponding genetic model estimated by Cox model analysis.
§frequency estimated in sample A.
¶p > 0.2.

Table 4 Results in the whole sample for the four *IFNGR2* variants providing the strongest evidence for association

Marker (type)	Position	Genotypes	No fibrosis F0–1, n (%)	Fibrosis F3–4, n (%)	Case/control study OR (95% CI)	p	Survival analysis HR (95% CI)	p
rs9976971 (G/A SNP)	33689967* 5′ region†	GG and AG	231 (89.2)	98 (73.7)	1		1	
		AA	28 (10.8)	35 (26.3)	2.95 (1.70 to 5.11)	8×10^{-5}	2.62 (1.76 to 3.91)	9×10^{-7}
		Total	259	133				
rs10600672 (AA/– ins/del)	33691495 5′ region	–/– and AA/–	228 (88.4)	98 (73.1)	1		1	
		AA/AA	30 (11.6)	36 (26.9)	2.79 (1.63 to 4.79)	10^{-4}	2.47 (1.66 to 3.66)	4×10^{-6}
		Total	258	134				
rs17882748 (T/C SNP)	33697591 Exon 1, 5′UTR	TT and CT	216 (83.7)	92 (68.7)	1		1	
		CC	42 (16.3)	42 (31.3)	2.35 (1.43 to 3.84)	6×10^{-4}	2.05 (1.40 to 3.01)	2×10^{-4}
		Total	258	134				
rs2284553 (G/A SNP)	33698565 Intron 1	GG and AG	233 (90.3)	105 (78.4)	1		1	
		AA	25 (9.7)	29 (21.6)	2.57 (1.44 to 4.61)	0.001	2.39 (1.57 to 3.66)	3×10^{-5}
		Total	258	134				

*Position in base pairs on chromosome 21.
†position relative to *IFNGR2* gene.

sequencing in 23 European subjects for a region of ~ 38 kb (33 695 200–33 733 200) (http://pga.gs.washington.edu/data/ifngr2/, accessed 5 May 2010). Finally, as rs9976971 is located in 5′ of *IFNGR2* (33 689 967), we also sequenced a sample of 32 French Caucasian subjects for a region of 12.3 kb from 33 689 894 to 33 702 179 (figure 1 and supplementary table 3). Based on these sequencing data from both the SeattleSNPs database and our own results, we identified two additional variants in strong LD with rs9976971 (table 4). One is the T/C SNP rs17882748 (r^2=0.83 with rs9976971) located in the untranslated region of exon 1, and the other is an AA insertion/deletion denoted as rs10600672 (r^2=0.98 with rs9976971) located in the 5′ region of the gene at position 33 691 495.

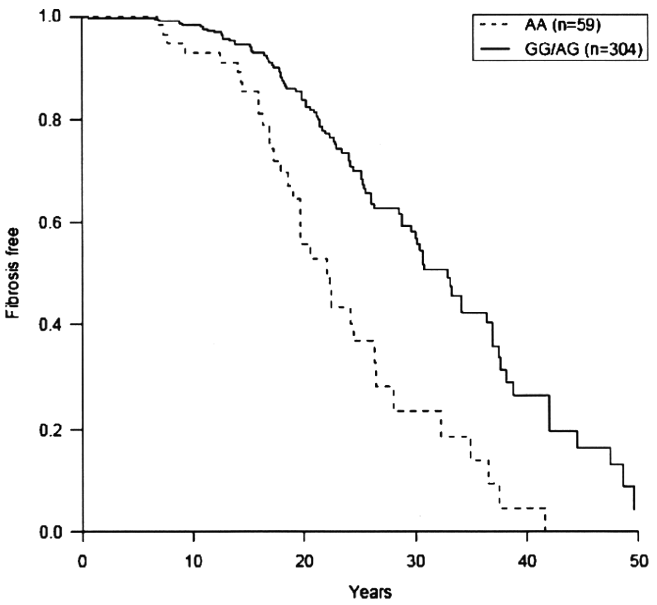

Figure 2 Effect of single nucleotide polymorphism rs9976971 on fibrosis progression. The figure shows the variation with time of the proportion of fibrosis-free patients according to genotypes at rs9976971. The time of follow-up was estimated from the presumed year of infection to the year of either the first biopsy showing severe fibrosis (F3–4 patients) or the last biopsy showing no fibrosis without any treatment (for F0–1 patients).

A cluster of four *IFNGR2* variants is strongly associated with liver fibrosis progression

Table 4 shows the results of both the case–control and the survival analysis with the four variants of interest over the combined samples A and B. Although all variants were strongly associated with development of severe fibrosis, the most significant results were observed with rs9976971 ($p=9\times10^{-7}$) and the AA ins/del ($p=4\times10^{-6}$) when using survival analysis. The AA insertion is in almost perfect LD with the A allele of rs9976971 (only three subjects had discordant genotypes) so that the HR of progressing towards severe fibrosis for AA/AA subjects as compared with AA/– or –/– subjects was 2.47 (1.66 to 3.66), and the curve of progression towards fibrosis with age according to this ins/del variant was extremely similar to that shown in figure 2 for rs9976971. We did not find any significant heterogeneity of these associations according to gender, mode of infection (blood transfusion/IVD use/others), viral genotypes (1 and 4 vs others), age at infection (≤20 years vs >20 years). We also conducted an analysis considering the different haplotypes that could be derived from these four variants using the method developed in the THESIAS program.[22] As expected by LD, two common haplotypes accounted for >90% of the estimated haplotypes (table 5). The common haplotype carrying the risk alleles at the four *IFNGR2* variants had a frequency of 0.338 and 0.447 in F0–1 and F3–4 patients, respectively. Under a recessive model (table 5), the effects of this at-risk haplotype (using a case–control or a survival analysis) were slightly lower than those estimated from rs9976971 or rs10600672 alone. As expected by the r^2 value at 0.98, the two variant haplotypes consisting of rs9976971 and rs10600672 provided results almost identical to the analysis of any of these variants alone (data not shown). However, none of the tested haplotypes consisting of three or four of these *IFNGR2* variants provided stronger evidence for association than rs9976971 or rs10600672 when considered alone.

DISCUSSION

In this study, we investigated whether the development of liver fibrosis in HCV chronically infected patients might be influenced by polymorphisms located in a panel of 36 genes involved in the fibrogenesis/fibrolysis process. To reduce the variability in assessing fibrosis stage by biopsy,[24] only F0–1 and F3–4 stages were included to define phenotypes. We found a single convincing signal of association in *IFNGR2*, which was the most

Table 5 Results of the analysis considering haplotypes derived from the cluster of the four *IFNGR2* variants providing the strongest evidence for association

Polymorphisms				Haplotype frequencies		Case–control study		Survival analysis	
rs9976971	rs10600672	rs17882748	rs2284553	No fibrosis	Fibrosis	OR (95% CI)	p	HR (95% CI)	p
G	–	T	G	0.570	0.466				
A	AA	C	A	0.338	0.447	2.70 (1.48 to 4.92)*	0.0011	2.32 (1.52 to 3.56)*	10^{-4}
G	–	T	A	0.049	0.034				
A	AA	T	A	0.035	0.034				
A	AA	C	G	0.004	0.004				
G	AA	C	A	0.004	0.004				

*Results obtained with a recessive model for the common haplotype A/AA/C/A consisting of the risk alleles of the four *IFNGR2* variants.

significant in our primary sample and was exactly replicated (same allele at risk and same genetic model) in our second sample. The evidence for association was even higher (by one order of magnitude) when using the information obtained by the time of progression, providing strong additional support for our findings. The *IFNGR2* signal results from a cluster of four variants in strong LD. These variants are quite common with a frequency in the HapMap Caucasian population of 0.41 for risk allele A of the most associated SNP rs9976971, and 26% AA homozygosity in our sample of F3–4 patients. Sequencing data obtained either from existing databases or from our present analysis excluded the role of any other SNPs located within a region of ~43.5 kb (33 689 894–33 733 200 bps) encompassing the *IFNGR2* gene. Analysis of the HapMap database also made quite unlikely the hypothesis that this signal could be due to another SNP in long-range LD with this cluster of four *IFNGR2* variants. Refined analysis showed that no haplotypes derived from this cluster of four variants provided stronger evidence for association than any of the four SNPs when analysed alone. From a statistical point of view, the strongest evidence was obtained with rs9976971 and the AA ins/del (rs10600672), which are in almost perfect LD. Nevertheless, the roles of rs2284553 (in intron 1) and rs17882748 (in untranslated region of exon 1) could not be ruled out.

Although several association studies have investigated the role of a small number of polymorphisms within the gene encoding IFNγ (*IFNG*) and progression to fibrosis in HCV infection, in particular a variant at position +874 which may influence IFNγ expression, no consistent and clearly replicated results have been reported.[10] A study also tested the role of variants within the interferon γ receptor 1 gene (*IFNGR1*), without any significant results.[25] In this context, it is interesting to note that two linkage studies conducted in Sudan[26] and Egypt[27] mapped a locus predisposing to development of severe liver fibrosis due to the parasite *Schistosoma mansoni* infection in a region including the *IFNGR1* gene. No precise variants underlying these linkage peaks have been reported yet. To our knowledge, no studies have yet investigated the influence of *IFNGR2* polymorphisms in HCV-related liver fibrosis. We cannot rule out the possibility that some *IFGNR2* variants were included in the study that tested the association with 24 823 putative functional SNPs.[12] However, it is unlikely that any of our four associated polymorphisms were included in that panel as they are not known to be functional.

IFNγ is usually considered to be an anti-fibrogenic cytokine. In an experimental model of liver fibrosis induced with carbon tetrachloride, IFNγ-deficient mice exhibited more pronounced hepatic fibrosis lesions than wild-type animals, and exogenous IFNγ administered to deficient animals reduced the level of fibrosis.[28] In human cells, IFNγ has been shown to inhibit activation, proliferation and collagen synthesis in cultures of activated hepatic stellate cells and hepatic myofibroblasts.[16 29] It is also interesting to note that IFNγ has been proposed as a treatment for idiopathic pulmonary fibrosis and was associated with reduced mortality in a meta-analysis.[30] Along the same lines, high levels of IFNγ production were associated with protection against periportal fibrosis in subjects infected with the parasites *S mansoni*[31] or *Schistosoma japonicum*.[32]

However, the role of IFNγ may be more complex, as suggested by a study performed on liver biopsy specimens from patients with chronic HCV infection, which showed that increased IFNγ expression was associated with portal inflammation and fibrosis stage.[33] This latter observation suggests that the proinflammatory effects of IFNγ may predominate over its anti-fibrogenic role in HCV liver fibrosis, although IFNγ expression might also be a consequence of the fibrosis process. For example, an impaired function of IFNγ receptors might lead to increased production of IFNγ, as seen in patients with complete IFNγ receptor deficiencies.[34]

In conclusion we have found that progression to severe liver fibrosis in HCV chronically infected patients of European origin is strongly associated with a cluster of four *IFNGR2* variants. Interestingly, between the two IFNγ receptors, IFNγR2 expression appears to be the deciding factor that controls the way in which target cells physiologically respond to IFNγ.[35 36] Functional studies are continuing in human liver cells to investigate the detailed biological mechanisms of this association and the potential effect of this cluster of variants on *IFNGR2* regulation. Prospective studies using large samples will also assess the predictive value of these polymorphisms with repeated validated non-invasive biomarkers.[37] IFNγ has already been found to be associated with clearance of HCV infection.[38 39] These results highlight the role of the IFNγ pathway in development of liver fibrosis that may pave the way for new treatments.

Author affiliations
[1]Unité d'Hépatologie, Institut National de la Santé et de la Recherche Médicale, U567, Hôpital Cochin, Paris, France
[2]Centre National de Génotypage, Evry, France
[3]Laboratory of Human Genetics of Infectious Diseases, Necker Branch, Institut National de la Santé et de la Recherche Médicale, U550, Necker Medical School, Paris, France
[4]Université Paris Descartes, Paris, France
[5]Center for the Study of Hepatitis C, The Rockefeller University, New York, USA
[6]Assistance Publique-Hôpitaux de Paris (AP-HP)/Université Pierre et Marie Curie (UPMC) Liver Center, Hôpital Pitié-Salpêtrière, Paris, France
[7]Banque de Cellules Cochin-APHP, Service de Biochimie et Génétique Moléculaire, Hôpital Cochin, Paris, France
[8]Department of Hepato-gastroenterology, Hôpital Saint Joseph, Marseille, France
[9]Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, The Rockefeller University, New York, USA
[10]Pediatric Hematology and Immunology Unit, Necker Children's Hospital, Paris, France
[11]Institut National de la Santé et de la Recherche Médicale, U785, Villejuif, France
[12]Unit of Human Disease Genomics, Institut National de la Santé et de la Recherche Médicale, Kyoto University, Kyoto, Japan

Acknowledgements We thank Rachel Morra, AP-HP/UPMC liver center, Hôpital Pitié-Salpêtrière, Paris, France, and la Banque de Cellules - Cochin APHP with all its staff for their logistic help and for providing patient lymphoblastoid cell lines. We are grateful to David-Alexandre Tregouët for helpful discussions about the haplotype analysis, and to the Centre National de Génotypage (CNG) for its help in genotyping. We also thank all members of the laboratory of Human Genetics of Infectious Diseases for fruitful discussions.

Funding This work was supported by grants from Institut National de la Santé et de la Recherche Médicale (4CH08G), Japanese Society for Promotion of Science (JSPS)/Inserm cooperation agreement, and Agence Nationale de Recherche sur le SIDA et les hépatites virales (ANRS HC EP13). LA was supported, in part, by a grant from Assistance Publique-Hôpitaux de Paris.

Competing interests None.

Ethics approval This study was conducted with the approval of the institutional review board (CPP: Comité de Protection des Personnes) of Ile de France - Paris - Saint Antoine, on 5 March 2002.

Contributors All authors participated in the collection, the management, and the interpretation of clinical data, and in the writing and the final approval of the manuscript. BN, ML, CB, TP, FM, SP and LA designed the study. SP, BR, EP, SH and LA were involved in statistical analysis. RLM and FM were involved in the sequencing of IFNGR2.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Lauer GM,** Walker BD. Hepatitis C virus infection. *N Engl J Med* 2001;**345**:41–52.
2. **Shepard CW,** Finelli L, Alter MJ. Global epidemiology of hepatitis C virus infection. *Lancet Infect Dis* 2005;**5**:558–67.
3. **Dustin LB,** Rice CM. Flying under the radar: the immunobiology of hepatitis C. *Annu Rev Immunol* 2007;**25**:71–99.
4. **Marcellin P,** Asselah T, Boyer N. Fibrosis and disease progression in hepatitis C. *Hepatology* 2002;**36**:S47–56.
5. **Missiha SB,** Ostrowski M, Heathcote EJ. Disease progression in chronic hepatitis C: modifiable and nonmodifiable factors. *Gastroenterology* 2008;**134**:1699–714.
6. **Poynard T,** Bedossa P, Opolon P. Natural history of liver fibrosis progression in patients with chronic hepatitis C. The OBSVIRC, METAVIR, CLINIVIR, and DOSVIRC groups. *Lancet* 1997;**349**:825–32.
7. **Poynard T,** Ratziu V, Charlotte F, *et al*. Rates and risk factors of liver fibrosis progression in patients with chronic hepatitis c. *J Hepatol* 2001;**34**:730–9.
8. **Ortiz V,** Berenguer M, Rayon JM, *et al*. Contribution of obesity to hepatitis C-related fibrosis progression. *Am J Gastroenterol* 2002;**97**:2408–14.
9. **Pol S,** Fontaine H, Carnot F, *et al*. Predictive factors for development of cirrhosis in parenterally acquired chronic hepatitis C: a comparison between immunocompetent and immunocompromised patients. *J Hepatol* 1998;**29**:12–19.
10. **Osterreicher CH,** Stickel F, Brenner DA. Genomics of liver fibrosis and cirrhosis. *Semin Liver Dis* 2007;**27**:28–43.
11. **Bataller R,** North KE, Brenner DA. Genetic polymorphisms and the progression of liver fibrosis: a critical appraisal. *Hepatology* 2003;**37**:493–503.
12. **Huang H,** Shiffman ML, Cheung RC, *et al*. Identification of two gene variants associated with risk of advanced fibrosis in patients with chronic hepatitis C. *Gastroenterology* 2006;**130**:1679–87.
13. **Huang H,** Shiffman ML, Friedman S, *et al*. A 7 gene signature identifies the risk of developing cirrhosis in patients with chronic hepatitis C. *Hepatology* 2007;**46**:297–306.
14. **Bataller R,** Brenner DA. Liver fibrosis. *J Clin Invest* 2005;**115**:209–18.
15. **Friedman SL.** Molecular regulation of hepatic fibrosis, an integrated cellular response to tissue injury. *J Biol Chem* 2000;**275**:2247–50.
16. **Lotersztajn S,** Julien B, Teixeira-Clerc F, *et al*. Hepatic fibrosis: molecular mechanisms and drug targets. *Annu Rev Pharmacol Toxicol* 2004;**45**:605–28.
17. **Bedossa P,** Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology* 1996;**24**:289–93.
18. **Fan JB,** Oliphant A, Shen R, *et al*. Highly parallel SNP genotyping. *Cold Spring Harb Symp Quant Biol* 2003;**68**:69–78.
19. **Dizier MH,** Bouzigon E, Guilloud-Bataille M, *et al*. Genome screen in the French EGEA study: detection of linked regions shared or not shared by allergic rhinitis and asthma. *Genes Immun* 2005;**6**:95–102.
20. **Takahashi M,** Matsuda F, Margetic N, *et al*. Automated identification of single nucleotide polymorphisms from sequencing data. *J Bioinform Comput Biol* 2003;**1**:253–65.
21. **Barrett JC,** Fry B, Maller J, *et al*. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;**21**:263–5.
22. **Tregouet DA,** Garelle V. A new JAVA interface implementation of THESIAS: testing haplotype effects in association studies. *Bioinformatics* 2007;**23**:1038–9.
23. **Hnatyszyn HJ.** Chronic hepatitis C and genotyping: the clinical significance of determining HCV genotypes. *Antivir Ther* 2005;**10**:1–11.
24. **Bedossa P,** Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 2003;**38**:1449–57.
25. **Falleti E,** Fabris C, Toniutto P, *et al*. Genetic polymorphisms of inflammatory cytokines and liver fibrosis progression due to recurrent hepatitis C. *J Interferon Cytokine Res* 2007;**27**:239–46.
26. **Dessein AJ,** Hillaire D, Elwali NE, *et al*. Severe hepatic fibrosis in Schistosoma mansoni infection is controlled by a major locus that is closely linked to the interferon-gamma receptor gene. *Am J Hum Genet* 1999;**65**:709–21.
27. **Blanton RE,** Salam EA, Ehsan A, *et al*. Schistosomal hepatic fibrosis and the interferon gamma receptor: a linkage analysis using single-nucleotide polymorphic markers. *Eur J Hum Genet* 2005;**13**:660–8.
28. **Shi Z,** Wakil AE, Rockey DC. Strain-specific differences in mouse hepatic wound healing are mediated by divergent T helper cytokine responses. *Proc Natl Acad Sci U S A* 1997;**94**:10663–8.
29. **Mallat A,** Preaux AM, Blazejewski S, *et al*. Interferon alfa and gamma inhibit proliferation and collagen synthesis of human Ito cells in culture. *Hepatology* 1995;**21**:1003–10.
30. **Bajwa EK,** Ayas NT, Schulzer M, *et al*. Interferon-gamma1b therapy in idiopathic pulmonary fibrosis: a metaanalysis. *Chest* 2005;**128**:203–6.
31. **Henri S,** Chevillard C, Mergani A, *et al*. Cytokine regulation of periportal fibrosis in humans infected with Schistosoma mansoni: IFN-gamma is associated with protection against fibrosis and TNF-alpha with aggravation of disease. *J Immunol* 2002;**169**:929–36.
32. **Arnaud V,** Li J, Wang Y, *et al*. Regulatory role of interleukin-10 and interferon-gamma in severe hepatic central and peripheral fibrosis in humans infected with Schistosoma japonicum. *J Infect Dis* 2008;**198**:418–26.
33. **Napoli J,** Bishop GA, McGuinness PH, *et al*. Progressive liver injury in chronic hepatitis C infection correlates with increased intrahepatic expression of Th1-associated cytokines. *Hepatology* 1996;**24**:759–65.
34. **Fieschi C,** Dupuis S, Picard C, *et al*. High levels of interferon gamma in the plasma of children with complete interferon gamma receptor deficiency. *Pediatrics* 2001;**107**:E48.
35. **Bach EA,** Szabo SJ, Dighe AS, *et al*. Ligand-induced autoregulation of IFN-gamma receptor beta chain expression in T helper cell subsets. *Science* 1995;**270**:1215–18.
36. **Bernabei P,** Coccia EM, Rigamonti L, *et al*. Interferon-gamma receptor 2 expression as the deciding factor in human T, B, and myeloid cell proliferation or death. *J Leukoc Biol* 2001;**70**:950–60.
37. **Poynard T,** Muntenau M, Morra R, *et al*. Methodological aspects of the interpretation of non-invasive biomarkers of liver fibrosis: a 2008 update. *Gastroenterol Clin Biol* 2008;**32**:8–21.
38. **Guidotti LG,** Chisari FV. Immunobiology and pathogenesis of viral hepatitis. *Annu Rev Pathol* 2006;**1**:23–61.
39. **Thimme R,** Oldach D, Chang KM, *et al*. Determinants of viral clearance and persistence during acute hepatitis C virus infection. *J Exp Med* 2001;**194**:1395–406.

Interferon γ receptor 2 gene variants are associated with liver fibrosis in patients with chronic hepatitis C infection

Bertrand Nalpas, Roubila Lavialle-Meziani, Sabine Plancoulaine, et al.

Gut 2010 59: 1120-1126 originally published online June 29, 2010
doi: 10.1136/gut.2009.202267