ARTICLE

Schistosomal hepatic fibrosis and the interferon gamma receptor: a linkage analysis using single-nucleotide polymorphic markers

Ronald E Blanton*,[1], Ekram Abdel Salam[2], Amgad Ehsan[3], Charles H King[1] and Katrina AB Goddard[4]

[1]Center for Global Health and Diseases, Case Western Reserve University, Wolstein Research Building, Room 4133, 2103 Cornell Road, Cleveland, OH 44106-7286, USA; [2]Department of Pediatrics, University of Cairo, Egypt; [3]Department of Nephrology, Division of Genetic Epidemiology, University of Cairo, Egypt; [4]Department of Epidemiology and Biostatistics, Case University, Cleveland, OH, USA

A minority of individuals infected with the parasite Schistosoma mansoni develops hepatic fibrosis. HLA studies in Egypt and a candidate gene search in a Sudanese population indicate that the host's genetics contribute to disease susceptibility. In an Egyptian community, 32.7% of individuals 11 years and older had significant fibrosis by WHO ultrasound criteria. Linkage to 10 candidate genes was tested using 89 affected sibling pairs from 40 pedigrees in this community. The candidates included genes that initiate fibrosis, participate in collagen synthesis, or control collagen degradation. Two to four single-nucleotide polymorphisms (SNPs) were genotyped per locus, and 188 individuals were genotyped at 48 markers. Model-free modified Haseman–Elston analysis identified linkage to a SNP in the interferon gamma receptor locus ($P = 0.000001$). There was also weak evidence for linkage to the interleukin 13-4 region and tissue growth factor beta 1.

Introduction

Only 5–10% of individuals infected with the blood fluke Schistosoma mansoni will develop hepatic fibrosis, and in population-based studies, intensity of infection does not correlate well with the presence of fibrosis by ultrasound. Thus, infection in and of itself is not equivalent to disease. This observation is central to understanding the distribution of many infectious diseases. For schistosomal fibrosis, different geographic regions display great variation in the severity of disease that is not well explained by the intensity or prevalence of infection alone.[1] This variability is also found between individuals living in the same endemic area. There is increasing evidence that, among other factors, host genetic differences account for a portion of this variability.[2] Dessein et al[2] conducted a search for four candidate genes related to S. mansoni induced hepatic fibrosis and found linkage in a region near the interferon gamma receptor 1 (IFNGR1) gene. However, this study was strongly biased in favor of examining areas directly related to immune response and not the final pathway of collagen deposition and maturation. In addition, even for genes that influence immune responses, additional good candidates exist.

The purpose of this work was to search for genes linked to hepatic fibrosis from among 11 candidates and in a

*Correspondence: Dr R Blanton, Center for Global Health and Diseases, Case Western Reserve University, Wolstein Research Building, Room 4133, 2103 Cornell Road, Cleveland, OH 44106-7286, USA.
Tel: +1 216 368 4818; Fax: +1 216 368 4825;
E-mail: reb6@po.cwru.edu

different population than the previous study. This study extends previous observations by reexamining linkage between schistosomal fibrosis and the already identified IFNGR1 region, as well as important candidates not assessed previously. We therefore chose to investigate genes regulating immunologic responses (IL10, IL13–IL4, IFNGR1, TGFβ1), genes regulating collagen production and stability (CTGF, Col1a1, Col3a1, LOX) and genes responsible for collagen degradation (MMP1, MMP2, TIMP1) Table 1. In addition, this work focused on multiple nuclear families in a population not previously studied.

Methods, subjects and materials
Population
Individuals over 11 years of age from a community in Kafr el Sheikh governorate, Nile Delta, Egypt were examined by portable ultrasound for evidence of schistosomal hepatic fibrosis. This region has traditionally had the country's highest prevalence of S. mansoni infection.[3] Approximately 70% of those eligible in the central community and some individuals from the nearby hamlets (ezbas) were surveyed ($n = 2051$). Blood was collected from first-degree relatives in families with more than one sibling with grade C or higher fibrosis (WHO criteria[4]). Prior to the study, there were three consecutive years of mass treatment for school children in this community followed by three more years of targeted chemotherapy for infected adults and children.

This study was performed in accordance with the ethical guidelines for human research set forth in the Declaration of Helsinki, under a protocol approved by the Human Investigation Review Boards of University Hospitals of Cleveland, OH, USA and Cairo University, Egypt. All participants gave free and informed consent for use of their blood for serologic, biochemical and genetic studies.

Phenotype determination and construction
The diagnosis and grading of schistosomal hepatic fibrosis was based on WHO criteria established in the year 2000 and modified for this study.[4,5] The standard consists of a qualitative determination of the overall condition of the liver graded according to predetermined patterns (image pattern). The standard also includes a quantitative measurement of the two-wall thickness of the first intrahepatic branch point of the portal vein (PBWT) and the diameter of the main portal vein (PVD). Magnification was used to increase accuracy of these measurements. The images were obtained by two Egyptian ultrasonographers and interpreted by a third radiologist not involved in the Egyptian study. The radiologists' agreement on image pattern was 84% (κ statistic $= 0.534$ where a statistic > 0.40 is considered good concurrence). The qualitative score or 'image pattern' was highly correlated with both quantitative measurements,[5,6] and the ultrasound examination has been validated histologically for sensitivity and specificity for hepatic fibrosis due to S. mansoni.[7,8]

Schistosomal fibrosis was modeled as both binary and quantitative traits based on ultrasound image pattern, PBWT and PVD. The binary phenotype was further defined as simple or severe fibrosis in order to allow comparisons with the only other linkage study for this condition. Simple fibrosis was defined as any image pattern whose severity was $\geq$ grade C. Severe fibrosis was defined $\geq$ as grade D. The severe fibrosis definition also included those with grade C patterns where the portal vein diameter was > 2 SD from the mean of those with normal image patterns (ie > 14.0 mm).

Quantitative traits consisted of the image pattern, the PBWT and PVD. The image pattern was converted to a quantitative scale by assigning each letter value (A–F) to a sequential integer (0–5). This scaling of the image pattern is linearly related to the continuous measurements PBWT and PVD.[6] Probands were considered affected if their image pattern was grade C or more severe. Families entered the study based on having at least two affected siblings. Although hepatitis B and C are common in this community, the hepatitis serology is known not to be associated with schistosomal fibrosis and was not, therefore, considered as a covariate in the analysis.[6]

Blood, plasma and DNA
In total, 10 ml of blood was collected from those with significant schistosomal fibrosis and their first-degree relatives. Plasma and buffy coat were separated and frozen at -20°C. DNA was isolated from the buffy coat using the Blood and Cell Culture Kit from Qiagen (Qiagen Inc., Valencia, CA, USA) according to the manufacturer's recommendation. Each sample was measured for DNA content by UV absorption spectrophotometry.

Candidate gene and marker selection
The candidates single-nucleotide polymorphisms (SNPs) were selected based on the weight of published evidence, their allele frequency, if known, and relative distance from other markers. For each gene, 4–6 SNPs were selected for genotyping. TNFα polymorphisms were excluded, due to difficulty identifying unique regions in the genome for primer design near the desired SNPs. Further, several studies have indicated no linkage[2] or association[9] with this gene and schistosomal fibrosis (S. mansoni and S. japonicum infection, respectively). Markers were selected within the following genes: interleukin 10 (IL10), interleukin 13 and 4 (IL13–IL4), IFNGR1, tissue growth factor beta 1 (TGFβ1), connective tissue growth factor (CTGF), type I collagen (Col1a1), type III collagen (Col3a1), lysyl oxidase (LOX), metalloproteinase 1 (MMP1), metalloproteinase 2 (MMP2), and tissue inhibitor of metalloproteinase 1 (TIMP1). DNAprint Genomics (Sarasota, FL, USA) performed all of the genotyping using the SNP-It (Orchid

Table 1 SNP selection

Chromosomal location	Gene	GenBank ID	Polymorphism		Validated	Structural location
			Base	Amino acid		
Immunologic factors						
1q31–q32	IL10	2222202	C/T	—	Y	Intron
1q31–q32	IL10	1554286	C/T	—	Y	Exon/intron
1q31–q32	IL10	3024496	C/T	—	Y	Exon
1q31–q32	IL10	3024500	C/T	—	Y	3′UTR
5q31–q31.1	IL13	rs2069739	A/G	—	Y	Promoter
5q31–q31.1	IL13	rs20541	C/T	Arg130Gln	Y	Coding
5q31–q31.1	IL13	rs2243202	C/T	—	Y	3′UTR & CNS
5q31–q31.1	IL4	rs2243250	C/T	—	Y	Promoter
6q23–q24	INFγ R1	rs608914	C/T	—	N	Promoter
6q23–q24	INFγ R1	rs1327474	A/G	—	N	Promoter
6q23–q24	INFγ R1	rs1327475	T/C	—	Y	First intron
6q23–q24	INFγ R1	rs1927050	C/T	—	N	First intron
6q23–q24	INFγ R1	rs3799488	C/T	—	N	Exon/intron
6q23–q24	INFγ R1	rs3799489	A/G	—	Y	Intron
19q13.1–13.3	TGFβ1	rs3826714	C/T	—	Y	3′UTR
19q13.1–13.3	TGFβ1	rs1056854	A/G	—	Y	3′UTR
19q13.1–13.3	TGFβ1	rs1982073	C/T	P/L	Y	Exon, coding
19q13.1–13.3	TGFβ1	rs1054797	A/G	—	N	Promoter
Collagen production						
2q31	Col3α1	rs1516454	C/T	—	Y	Intron
2q31	Col3α1	rs2271681	C/T	—	Y	Intron/exon
2q31	Col3α1	rs1800255	A/G	Ala/Thr	Y	Exon, coding
2q31	Col3α1	rs3736490	G/A	—	Y	Intron
6q23	CTGF	rs2151532	G/A	—	Y	Promoter
6q23	CTGF	rs15375	A/G	—	N	Exon
6q23	CTGF	rs15960	A/G	—	N	3′ downstream
6q23	CTGF	rs11575	G/A	—	N	Not mapped
17q21.3–q22.1	Col 1α1	rs2249492	G/A	—	Y	Intron/exon
17q21.3–q22.1	Col 1α1	rs1007086	G/A	—	Y	Intron/exon
17q21.3–q22.1	Col 1α1	rs2734278	G/A	—	N	5′UTR
17q21.3–q22.1	Col 1α1	rs2696247	G/A	—	Y	Intron/exon
17q21.3–q22.1	Col 1α1	rs2277633	C/T	—	Y	Intron
Collagen stability						
5q23.2	LOX	rs3792803	A/G	—	Y	Intron
5q23.2	LOX	rs2434979	C/T	Gly/Gly	N	Exon coding
5q23.2	LOX	rs1800449	A/G	Arg/Glu	Y	Exon coding
5q23.2	LOX	rs3805443	A/G	—	Y	Intron
11q22	MMP1	rs470221	A/G	—	Y	Intron
11q22	MMP1	rs491152	C/T	—	Y	Intron/exon
11q22	MMP1	rs2075847	C/T	—	Y	Promoter
11q22	MMP1	rs514921	C/T	—	Y	Promoter
16q13–q21	MMP2	rs2287074	G/A	Thr/Thr	Y	Exon coding
16q13–q21	MMP2	rs243842	G/A	—	Y	Intron
16q13–q21	MMP2	rs2287076	C/T	—	Y	Intron/exon
16q13–q21	MMP2	rs14070	C/T	—	N	Not mapped
Xp11.3–p11.2	TIMP1	rs4898	C/T	—	N	Not mapped
Xp11.3–p11.2	TIMP1	rs1984392	A/G	—	N	Intron/exon
Xp11.3–p11.2	TIMP1	rs1043428	C/T	Pro/Pro	N	Exon coding
Xp11.3–p11.2	TIMP1	rs1062849	C/T	Ser/Phe	N	Exon coding
Xp11.3–p11.2	TIMP1	rs1803571	G/A	Ile/Val	N	Exon coding

IL – interleukin; IFNGR – interferon gamma receptor; TGF – transforming growth factor; Col 3α1 – type III collagen, CTGF – connective tissue growth factor, Col 1α1 – type I collagen; LOX – lysyl oxidase; MMP – metalloproteinase, TIMP – tissue inhibitor of metalloproteinase, CNS – conserved noncoding sequence.

Biosciences, Princeton, NJ, USA) ultrahigh throughput single-nucleotide extension method.[10] The program Genehunter (version 2.1)[11] was used to determine the information content of the markers.

Model-free linkage analysis
Linkage to schistosomal hepatic fibrosis was analyzed using the model-free, modified Haseman–Elston method[12] implemented in the program Sibpal (Statistical Analysis for

Genetic Epidemiology. S.A.G.E. v. 4.3, 2002. Statistical Solutions Ltd, Cork, Ireland). The allele frequency from the study sample was used for all markers. Linkage was assessed in both two point and multipoint analysis of the dichotomous and quantitative trait definitions. Distances between genes were provided by the Marshfield map in NCBI's Entrez Genome Mapview.

Since the PBWT and PVD were not collinear with the ultrasound image in regression analysis, they may contain distinct information about the phenotype. Further, age and sex have been shown to be significantly associated with fibrosis. The effect of these covariates on linkage was assessed with the program Lodpal (S.A.G.E. v. 4.3. Statistical Solutions Ltd, Cork, Ireland). Models were analyzed with and without the covariate or combinations of covariates and the differences in their Lod scores were determined.[13] Linkage disequilibrium between markers was tested by Fisher's exact test in the program Genepop (http://wbiomed.curtin.edu.au/genepop/).[14] Due to the relatively small number of markers used, no correction was made for consanguinity.

Results

Pedigree structure

A total of 40 pedigrees were collected that contained multiple affected siblings and whose members provided a blood sample (Table 2). There were 36 nuclear families and four larger families that included grandparents, uncles, aunts or cousins. There were no known marriage loops and two sets of dizygotic twins. The average family size considering only those individuals available for genotyping was 5.70 ± 1.65. There were 46 sibships with an average size of 3.09 ± 1.30. The mean age of the affected siblings under the simple definition of fibrosis was 35 ± 11.01 compared to 42 ± 12.07 for those affected in the general population.

Phenotype distribution

For the simple definition of fibrosis, the finding of C grade pathology was sufficient for identifying affecteds. Under the severe fibrosis phenotype, only 14% (7/48) of individuals with the C pattern were considered affected (Table 3). Of those who fulfilled the simple criteria of fibrosis, 58% were considered affected under the more severe definition. In the whole population, the prevalence of fibrosis under the simple definition was 36.8% and using severe fibrosis definition was 21.6%. There was no difference in age using the simple or severe fibrosis phenotypes (39.20 ± 16.08 and 39.80 ± 11.31, respectively, $P = 0.06$) or PBWT (4.51 ± 1.68 and 4.83 ± 1.64, respectively, $P = 0.25$). Mean PVD, however, was significantly smaller under the simple definition of the phenotype (13.73 ± 3.36 and 15.20 ± 2.97, respectively, $P < 0.01$). There was a higher male prevalence of both schistosome infection and hepatic fibrosis that increased from 2:1 for the simple fibrosis phenotype to 3:1 for the severe fibrosis phenotype. Using the simple fibrosis phenotype, there were 89 affected sib pairs and 100 affected relative pairs. With the severe fibrosis phenotype, there were 37 affected sib pairs and 42 affected relative pairs.

Table 2 Summary of pedigree characteristics

%Male	62.7%
Mean age	39.7 ± 14.3
Pedigrees (mean size[a])	40 (5.70 ± 1.65)
Nuclear families	36
Extended families	4
Sibships (mean size[a])	46 (3.09 ± 1.30)
Pairs	
Sis/Sis	16
Bro/Bro	109
Bro/Sis	62

[a]Includes only those who were genotyped.

Table 3 Summary of phenotypic characteristic

Simple fibrosis[a]		Severe fibrosis[b]	
Prevalence		Prevalence	
Population[c]	36.8%	Population[c]	21.6%
Sample[d]	51.8%	Sample[d]	27.2%
Mean age (sample)	39.13 ± 13.16 SD	Mean age (sample)	42.29 ± 15.25 SD
Male:female	>2:1	Male:female	>3:1
Affected sib pairs	89	Affected sib pairs	37
Affected relative pairs	100	Affected relative pairs	42

[a]Affected = C grade or more severe ultrasound pattern, $n = 110$.
[b]Affected = D grade or higher plus C grade with PVD > 14.8 mm, $n = 62$.
[c]Total population examined.
[d]Those entered in pedigrees analyzed for linkage.

Table 4 Informative marker characteristics

Chromosomal location	Gene	Locus ID (rs)	Hetero[a] (%)	Failed[b] (%)	HWP[c] (P-value)
1q31–q32	IL10	2222202	39.7	31.7	0.117
1q31–q32	IL10	1554286	28.4	0.0	0.903
1q31–q32	IL10	3024496	46.8	1.1	0.724
1q31–q32	IL10	3024500	44.3	2.7	0.342
2q31	Col3a1	1516454	22.8	0.5	0.740
2q31	Col3a1	2271681	2.7	1.6	0.980
2q31	Col3a1	3817316	37.2	1.1	0.322
2q31	Col3a1	3736490	40.0	2.7	0.908
5q23.2	LOX	3792803	1.1	1.6	0.992
5q23.2	LOX	1800449	19.9	2.2	0.826
5q31–q31.1	IL13–IL4	2069739	10.1	0.5	0.765
5q31–q31.1	IL13–IL4	20541	37.2	3.8	0.649
5q31–q31.1	IL13–IL4	2243202	5.4	2.7	0.931
5q31–q31.1	IL13–IL4	2243250	38.9	2.7	0.936
6q23	CTGF	2151532	14.3	0.5	0.568
6q23	CTGF	15375	0.5	1.6	0.515
6q23–q24	IFNGRI	608914	37.4	4.3	0.893
6q23–q24	IFNGRI	1327474	51.5	19.9	0.150
6q23–q24	IFNGRI	1327475	33.7	3.2	0.804
6q23–q24	IFNGR1	3799488	11.1	0.5	0.384
11q22	MMP1	2075847	12.8	1.6	0.604
11q22	MMP1	514921	36.7	1.1	0.570
11q22	MMP1	491152	2.7	3.2	0.981
11q22	MMP1	470221	35.1	2.7	0.015
16q13–q21	MMP2	2287076	41.2	1.6	0.497
16q13–q21	MMP2	243842	40.0	5.4	0.980
16q13–q21	MMP2	2287074	42.4	3.2	0.955
16q13–q21	MMP2	14070	50.8	4.8	0.629
17q21.3–q22.1	Col1a1	1007086	50.5	1.1	0.459
17q21.3–q22.1	Col1a1	2277633	0.5	0.5	0.992
17q21.3–q22.1	Col1a1	2249492	43.9	1.6	0.560
19q13.1–13.3	TGFβ1	3826714	42.8	1.6	0.538
19q13.1–13.3	TGFβ1	1056854	31.2	0.5	0.874
Xp11.3–p11.2	TMP1	1062849	1.1	2.2	0.992
Xp11.3–p11.2	TMP1	1043428	2.6	0.5	<0.0001

[a]Heterozygosity observed in the sample.
[b]Percent of samples that could not be genotyped.
[c]Hardy–Weinberg proportions assessed by goodness-of-fit test (P-value).

Genotyping

A total of 186 individuals were genotyped in 40 pedigrees at 48 SNP markers targeted to 11 genes. Of the 48 markers used for typing, 35 (72.9%) amplified successfully and were polymorphic (Table 4). Of these, 20% had minor allele frequencies <0.05 and correspondingly low heterozygosities and information content. The mean single point information content for the 35 polymorphic alleles was 0.15 ± 0.09. For the X-linked TIMP1, all markers for this gene either failed to amplify or were nearly monoallelic, so, linkage to this gene could not be further assessed.

Four markers had an individual failure rate in excess of 80% and nine were monomorphic. All individual DNA samples could be typed at most loci, so that DNA quality was not considered a factor in typing failures. The failures are likely to represent polymorphisms in the primer target sequence on the template. The dbSNP site of Genebank (www.ncbi.nlm.nih.gov/SNP/) provided allele frequencies for 62.0% of the selected SNPs. The expected crude average heterozygosity of successfully genotyped SNPs based on the dbSNP database was 29.6%, compared to 27.4% observed for these same SNPs in the Shamarka population.
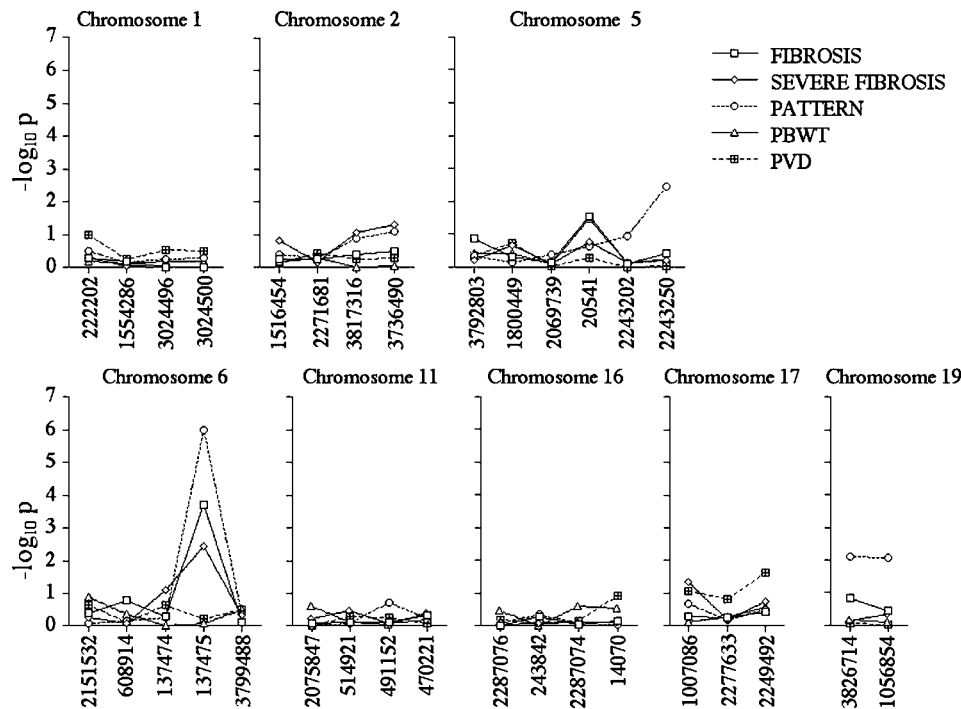

Figure 1 Converted *P*-value for single point modified Haseman–Elston linkage analysis. −Log$_{10}$ (*P*-value) for linkage to schistosomal fibrosis were plotted for each marker and for each definition of the phenotype. Significant linkage is indicated by a −Log$_{10}$ (*P*-value) = 2.34, that is, *P*-value of 0.0046 was considered significant after correction for multiple tests. Simple – simple definition of fibrosis used for phenotype; Severe – severe definition of fibrosis used for phenotype; PATTERN – ultrasound image pattern; PBWT – portal branch wall thickness; PVD – portal vein diameter. *P*-values available in supplemental Table 5.

Single point analysis

There were 11 candidate genes tested for linkage. The markers within each gene were not considered independent; therefore, the Bonferroni correction was applied for testing multiple genes, and a *P*-value of 0.0045 ($\alpha = 0.05/11$) was the chosen level of significance. With single point analysis the most significant *P*-value, 0.0000013, was obtained for the IFNGR1 marker rs1327475 (chromosome 6) using the ultrasound pattern as the quantitative trait (Figure 1). For the binary trait definition, there was significant evidence for linkage, using either the simple or the severe fibrosis phenotype ($P < 0.001$ and 0.004, respectively). Two additional markers in other genes produced significant *P*-values only when the ultrasound pattern was used. For the marker rs2243250 in the IL4 promoter (chromosome 5), the *P*-value was 0.004. Although the *P*-values were higher than the limit set by Bonferroni correction, both TGFβ1 markers (chromosome 14) had small *P*-values near the corrected level of significance. Other markers within the IFNGR1 gene or the IL13–IL4 gene clusters did not produce noticeably smaller *P*-values than markers in other regions.

Multipoint analysis

Five loci or regions (Col3a1, IL13–IL4, IFNGR1, Col1a1 and TGFβ1) were found to have *P*-values <0.05 (Figure 2). Only

the marker rs1056854 in the TGFβ1 gene had a score that was significant after correction for multiple tests, and only when the ultrasound pattern was used as the trait. Only the IFNGR1 had multipoint *P*-values of <0.05 for more than one phenotype. Not surprisingly, three of the four SNPs across the IFGR1 gene are in strong linkage disequilibrium (rs608914, rs1327474, rs1327475) and thus may not contribute much information to the multipoint analysis.

Effect of covariates

Neither age, sex, nor PBWT were found to affect linkage based on either the simple or severe fibrosis phenotype. Lodpal analysis did, however, show a significant effect of PVD for the IFNGR1 marker rs1327475 when the severe phenotype was used.

Discussion

In this study we used SNP markers to search for evidence of linkage to 11 candidate genes in a population with Egypt's highest prevalence of infection. We found evidence for linkage of schistosomal hepatic fibrosis to a polymorphism in the INFGR1 gene and suggestive evidence for linkage to a locus in the IL13–IL4 region and TGFβ1. The rs1327475 marker of IFNGR1 produced the most significant *P*-value (0.0000013) in model-free single point analysis. This was a
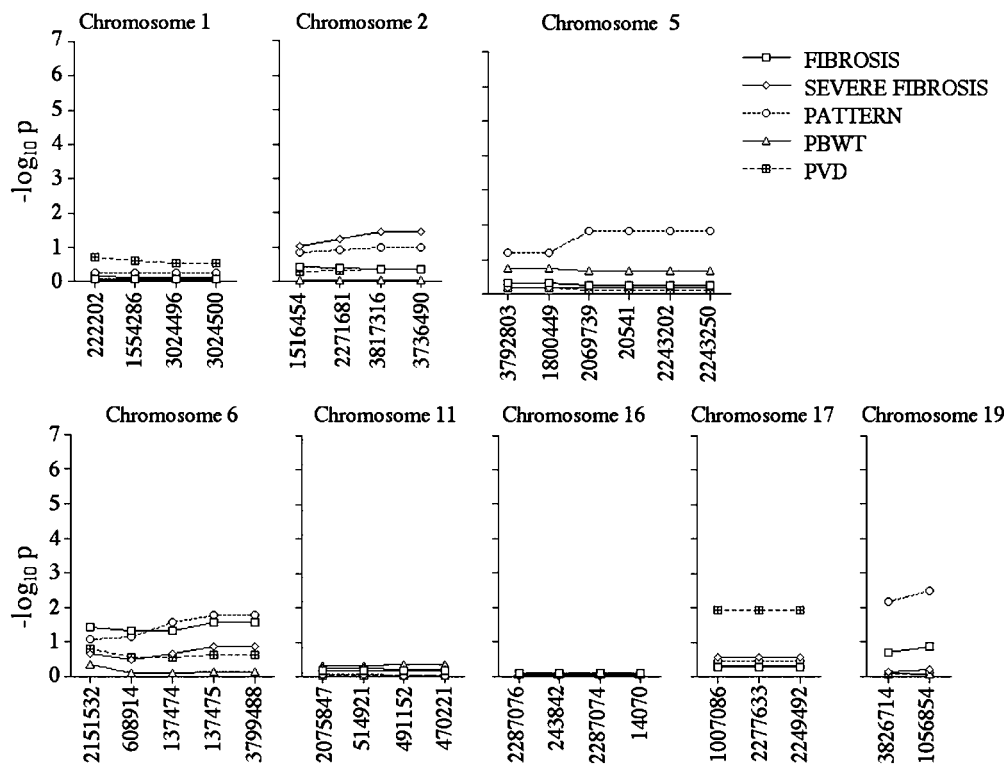
Figure 2 Converted *P*-value for multipoint modified Haseman–Elston linkage analysis. $-\text{Log}_{10}$ (*P*-value) for linkage to schistosomal fibrosis were plotted for each marker and for each definition of the phenotype. Significant linkage is indicated by a $-\text{Log}_{10}$ (*P*-value) $= 2.34$, that is, *P*-value of 0.0046 was considered significant after correction for multiple tests. Simple – simple definition of fibrosis used for phenotype; Severe – severe definition of fibrosis used for phenotype; PATTERN – ultrasound image pattern; PBWT – portal branch wall thickness; PVD – portal vein diameter. *P*-values available in supplemental Table 6.

significant value even after correction for multiple tests. In addition, this marker consistently showed significance under several phenotype definitions. Linkage to this gene is further supported by studies with IFNGR1 null mice that indicate this is a fibrosis-prone phenotype. Finally, the only other linkage study for this disease also identified region of the IFNGR1 gene as linked to the development of severe schistosomal fibrosis.[2]

γ interferon has long been associated with a protective affect against the development of fibrosis in *S. mansoni* infection[15] as well as liver cells undergoing other forms of damage.[16,17] The cytokine's receptor has only recently been the focus of study due primarily to the availability of IFNGR1 null mutant mice. Rezende *et al*[18] confirmed the lack of any effect on resistance to infection in these mice, but convincingly demonstrated both smaller granuloma size and rapid development of intense fibrosis in IFNGR1 null animals. In these animals, maturation of monocytes to macrophages failed or was delayed,[19] and migration of B and CD4 + cells into the granulomas was also delayed. The numbers of connective tissue cells and their *in vitro* proliferation, however, were increased.

IFNGR1 null mice also develop more hepatic fibrosis in response to intracellular infection with the murine gammaherpes virus.[20] This indicates that polymorphisms in this gene may be associated with a generalized susceptibility to hepatic fibrosis. This is particularly intriguing, given the more severe pathology found in this Egyptian population for both schistosomiasis and hepatitis C compared to a Kenyan population where both infections result in much more benign disease with respect to the liver.[1]

In humans, severe defects in IFNGR1 are probably selected against and are therefore rare.[21,22] There have been genetic studies, however, that associated common polymorphisms (allele frequency ∼0.5) in the IFNGR1 gene to presumably milder defects such as increased susceptibility to cerebral malaria or *Helicobacter pylori* infection.[23,24] In both cases, morbidity was associated with a promoter polymorphism (−56 C/T, or rs2234711). Reporter gene constructs with the 'C' mutation were found to produce lower levels of expression and perhaps explain differences in morbidity on this basis.[25] Unfortunately, this effect had not been reported at the time of this analysis and this marker was not included. The marker rs1327475,

however, is 4000 nucleotides downstream of rs2234711 and may be in linkage disequilibrium with this locus.

The strength of the evidence needs to be qualified somewhat since only one marker in the IFNGR1 gene showed significance and since multipoint analysis failed to increase significance at these markers. This may be due to the use of closely spaced SNP markers. In association studies, linkage disequilibrium between adjacent SNPs does not consistently fall or rise with distance when the interval is short (5–32 kb).[26–28] Saunders et al[29] also observed a wide variation in the strength of association between individual SNPs and disease even in simulated data. This is often explained by varying or low allele frequencies, but even some common polymorphisms near a functional locus may show no association. Linkage studies rely on information from recombination, but over short distances, the effects of mutation, genetic drift, and population admixture outweigh those of recombination.[30] One of the few studies to use SNPs for linkage analysis also found more evidence for linkage using single point analysis than multipoint linkage.[31] Since linkage analysis relies on recombination and currently available methods assume linkage equilibrium, a failure to increase significance in multipoint analysis may be in the nature of linkage studies with closely positioned SNP markers.[32–34] For closely spaced markers it appears that more than four good quality SNPs should be used for linkage or association studies.

Two additional genes had markers that were significant after correction (IL13–IL4) or had scores near this level (TGFβ1). The IL13–IL4 region produced a significant Lod score only for one marker, and only when the ultrasound pattern was used as the trait. Although the TGFβ1 markers were only near the level of significance, both markers that could be typed for this locus had similarly low P-values. This suggests they may have an effect on the phenotype, but are limited to a few families or are less influential than the INFGR1.

In addition to locus heterogeneity, demographic or environmental factors may affect the power to detect linkage. Within the population, fibrosis varied significantly by age and sex. The loci that demonstrated significance with single point analysis were assessed with respect to age, sex, PBWT and PVD to determine if these covariates influenced linkage. For the simple fibrosis phenotype, none of these was found to be significant. The age difference between sibs appears to be sufficiently small that it did not contribute much information for linkage, and the difference in rates of fibrosis by sex does not have a genetic origin at these loci. It is likely that the lower participation of women in agriculture and exposure to infection in this community accounts for much of the sex differences in fibrosis.

In order to permit comparisons with other studies, two definitions for fibrosis were employed. Under the latest WHO guidelines, the C ultrasound pattern of pathology is specific for schistosomal fibrosis, and so was used as the only ultrasound criterion for collecting families in Egypt. The only other linkage study for this disease to date defined fibrosis by the equivalent of the more severe D pattern and also included those with C pattern and PVD dilatation. Stratification by this severe phenotype only weakened the evidence for linkage by reducing the sample size. This suggests that the simple fibrosis phenotype is likely to have the same genetic basis as the more severe definition and uses more of the available information from the population.

The identification of genes linked to hepatic fibrosis is a starting point for investigation. This work in addition to providing strong support for IFNGR1 as a major gene locus in schistosomal fibrosis, suggests that genes in the region of IL13–IL4 and TGFβ1 are also linked. A limitation of this or any candidate gene approach is that out of thousands of genes, up to half of which are of unknown function, enough information is available that less than a dozen can be selected for testing. Schistosomal hepatic fibrosis still awaits evaluation with a genome-wide scan, the best available method to broadly survey for additional loci. It would also provide perspective on the relative importance of the loci identified here.

References

1 Blanton RE, Salam EA, Kariuki HC et al: Population-based differences in Schistosoma mansoni- and hepatitis C-induced disease. J Infect Dis 2002; **185**: 1644–1649.

2 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**: 709–721.

3 Cline BL, Richards FO, el Alamy MA et al: 1983 Nile Delta schistosomiasis survey: 48 years after Scott. Am J Trop Med Hyg 1989; **41**: 56–62.

4 Richter J, Hatz C, Campagne G, Berquist N, Jenkins J: Ultrasound in schistosomiasis. A practical Guide to the standardized use of ultrasonography for the assessment of schistosomiasis-related morbidity; Second International Workshop, Niamey, Niger, UNDP/World Bank/WHO Special Progamme for Research & Training in Tropical Diseases, 2000, pp 1–56.

5 Kariuki HC, Mbugua G, Magak P et al: Prevalence and familial aggregation of schistosomal liver morbidity in Kenya: evaluation by new ultrasound criteria. J Infect Dis 2001; **183**: 960–966.

6 King CH, Magak P, Salam EA, Ouma JH, Kariuki HC, Blanton RE: Measuring morbidity in Schistosomiasis mansoni: the relationship between image pattern, portal branch, and portal vein measurements in new WHO guidelines for ultrasound in schistosomiasis. Int J Trop Med 2002; **8**: 109–117.

7 Homeida M, Abdel-Gadir AF, Cheever AW, Bennett JL: Diagnosis of pathologically confirmed Symmer's periportal fibrosis by ultrasonography: a prospective blinded study. *Am J Trop Med Hyg* 1988; **38**: 86–91.

8 Abdel-Wahab M, Esmat G, Milad M, Abdel-Razek S, Strickland G: Characteristic sonographic pattern of schistosomal hepatic fibrosis. *Am J Trop Med Hyg* 1989; **40**: 72–76.

9 Hirayama K: Genetic factors associated with development of cerebral malaria and fibrotic schistosomiasis. *Korean J Parasitol* 2002; **40**: 165–172.

10 Bell PA, Chaturvedi S, Gelfand CA *et al*: SNPstream UHT: ultra-high throughput SNP genotyping for pharmacogenomics and drug discovery. *Biotechniques* 2002; **74** (Suppl 70–72): 76–77.

11 Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; **58**: 1347–1363.

12 Elston RC, Buxbaum S, Jacobs KB, Olson JM: Haseman and Elston revisited. *Genet Epidemiol* 2000; **19**: 1–17.

13 Goddard KA, Witte JS, Suarez BK, Catalona WJ, Olson JM: Model-free linkage analysis with covariates confirms linkage of prostate cancer to chromosomes 1 and 4. *Am J Hum Genet* 2001; **68**: 1197–1206.

14 Rousset F, Raymond M: Testing heterozygote excess and deficiency. *Genetics* 1995; **140**: 1413–1419.

15 Czaja MJ, Weiner FR, Takahashi S *et al*: Gamma-interferon treatment inhibits collagen deposition in murine schistosomiasis. *Hepatology* 1989; **10**: 795–800.

16 Rockey DC, Chung JJ: Interferon gamma inhibits lipocyte activation and extracellular matrix mRNA expression during experimental liver injury: implications for treatment of hepatic fibrosis. *J Invest Med* 1994; **42**: 660–670.

17 Baroni GS, D'Ambrosio L, Curto P *et al*: Interferon gamma decreases hepatic stellate cell activation and extracellular matrix deposition in rat liver fibrosis. *Hepatology* 1996; **23**: 1189–1199.

18 Rezende SA, Oliveira VR, Silva AM, Alves JB, Goes AM, Reis LF: Mice lacking the gamma interferon receptor have an impaired granulomatous reaction to *Schistosoma mansoni* infection. *Infect Immun* 1997; **65**: 3457–3461.

19 Oliveira VR, El-Cheikh MC, Aguiar AM *et al*: *Schistosoma mansoni* egg-induced hepatic granulomas in mice deficient for the interferon-gamma receptor have altered populations of macrophages, lymphocytes and connective tissue cells. *Microbes Infect* 2000; **2**: 1817–1826.

20 Ebrahimi B, Dutia BM, Brownstein DG, Nash AA: Murine gammaherpesvirus-68 infection causes multi-organ fibrosis and alters leukocyte trafficking in interferon-gamma receptor knock-out mice. *Am J Pathol* 2001; **158**: 2117–2125.

21 Jouanguy E, Lamhamedi-Cherradi S, Altare F *et al*: Partial interferon-gamma receptor 1 deficiency in a child with tuberculoid bacillus Calmette-Guerin infection and a sibling with clinical tuberculosis. *J Clin Invest* 1997; **100**: 2658–2664.

22 Altare F, Jouanguy E, Lamhamedi-Cherradi S *et al*: A causative relationship between mutant IFNgR1 alleles and impaired cellular response to IFNgamma in a compound heterozygous child. *Am J Hum Genet* 1998; **62**: 723–726.

23 Koch O, Awomoyi A, Usen S *et al*: IFNGR1 gene promoter polymorphisms and susceptibility to cerebral malaria. *J Infect Dis* 2002; **185**: 1684–1687.

24 Thye T, Burchard GD, Nilius M, Muller-Myhsok B, Horstmann RD: Genomewide linkage analysis identifies polymorphism in the human interferon-gamma receptor affecting *Helicobacter pylori* infection. *Am J Hum Genet* 2003; **72**: 448–453.

25 Juliger S, Bongartz M, Luty AJ, Kremsner PG, Kun JF: Functional analysis of a promoter variant of the gene encoding the interferon-gamma receptor chain I. *Immunogenetics* 2003; **54**: 675–680.

26 Mateu E, Calafell F, Lao O *et al*: Worldwide genetic analysis of the CFTR region. *Am J Hum Genet* 2001; **68**: 103–117.

27 Bourgain C, Genin E, Clerget-Darpoux F: Comparison of family based haplotype methods using intragenic SNPs in candidate genes. *Eur J Hum Genet* 2002; **10**: 313–319.

28 Bensen JT, Langefeld CD, Hawkins GA *et al*: Nucleotide variation, haplotype structure, and association with end-stage renal disease of the human interleukin-1 gene cluster. *Genomics* 2003; **82**: 194–217.

29 Saunders CL, Crockford GP, Bishop DT, Barrett JH: Using single nucleotide polymorphisms to investigate association between a candidate gene and disease. *Genet Epidemiol* 2001; **21** (Suppl 1): S415–S420.

30 Cox A, Camp NJ, Nicklin MJ, di Giovine FS, Duff GW: An analysis of linkage disequilibrium in the interleukin-1 gene cluster, using a novel grouping method for multiallelic markers. *Am J Hum Genet* 1998; **62**: 1180–1188.

31 Zhu X, Bouzekri N, Southam L *et al*: Linkage and association analysis of angiotensin I-converting enzyme (ACE)-gene polymorphisms with ACE concentration and blood pressure. *Am J Hum Genet* 2001; **68**: 1139–1148.

32 Abecasis GR, Wigginton JE: Linkage Analysis with markers that are in linkage disequilibrium; American Society of Human Genetics Annual Meeting, Toronto 2004, Abstract 93.

33 Huang Q, Shete S, Amos C: Ignoring linkage disequilibrium between markers induces false positive evidence of linkage for affected sib-pair analysis; American Society of Human Genetics Annual Meeting, Toronto 2004, Abstract 94.

34 Broman KW, Feingold E: SNPs made routine. *Nat Methods* 2004; **1**: 104–105.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)