

Localization of the Hemochromatosis Disease Gene: Linkage Disequilibrium Analysis Using an American Patient Collection

Submitted 02/08/96

(communicated by Ernest Beutler, M.D., 02/09/96)

Nicole K. Seese^{1,6}, Charles P. Venditti^{1,6}, Karen A. Chorney¹, Glenn S. Gerhard², Junli Ma³, Thomas J. Hudson³, Pradyumna D. Phatak⁴ and Michael J. Chorney^{1,5}

ABSTRACT: The genetic basis of idiopathic hemochromatosis, a common disorder of iron metabolism, has remained an enigma for over two decades. In an attempt to refine the chromosomal localization of this gene, we have conducted a linkage disequilibrium mapping study utilizing a large group of unrelated American patients. The 12 microsatellites used as genetic markers in this analysis include a series of recently described polymorphic dinucleotide (D6S1558, D6S1545 and D6S1554) and tetranucleotide (D6S1016 and D6S1281) repeats which map between D6S105 and D6S299. Haplotype reconstructions indicate that a core genotype, composed of D6S464 allele 3/D6S1260 allele 4/D6S1558 allele 5, exists on a majority of disease chromosomes. Stringent statistical measures of marker-disease disequilibrium suggest that only associations with D6S1260 are significant and furthermore, aid in the assignment of refined centromeric and telomeric limits for the likely location of the hemochromatosis gene. In summary, the genetic data presented in this report predict that the hemochromatosis locus resides between D6S464 and D6S1558, most likely very close to marker D6S1260. Because a single yeast artificial chromosome clone contains all three of the above loci, a thorough search for coding sequences in this region is likely to identify the gene mutated in this common disorder.

Keywords: hemochromatosis, linkage disequilibrium mapping, chromosome 6, MHC class I region

INTRODUCTION

Idiopathic hemochromatosis (HFE) [Mendelian Inheritance in Man (MIM, entry number 230800)] is an autosomal recessive disease characterized by insidious and

destructive iron overload. HFE is especially common in Caucasians of European ancestry and is one of the most common genetic diseases afflicting this population (1). Several studies designed to examine the incidence of hemochromatosis in geographically diverse

¹ Department of Microbiology and Immunology, The Pennsylvania State University College of Medicine, Hershey, PA 17033;

² Department of Pathology, The Pennsylvania State University College of Medicine, Hershey, PA 17033;

³ Whitehead Institute-Massachusetts Institute of Technology Center for Genome Research, Cambridge, MA 02639;

⁴ Mary M. Gooley Hemophilia Center and the Department of Medicine, The University of Rochester School of Medicine and Dentistry, Rochester, NY 14627;

⁵ Department of Pediatrics, The Pennsylvania State University College of Medicine, Hershey, PA 17033.

⁶ The contributions of the first and second authors can be considered equivalent.

Reprint request to: Michael J. Chorney, Ph.D., Department of Microbiology and Immunology, The Pennsylvania State University College of Medicine, The Milton S. Hershey Medical Center, Hershey, Pennsylvania 17033 USA, phone (717) 531-4604, fax (717) 531-3562, e-mail: mjcl18@psu.edu

©The Blood Cells Foundation, La Jolla, California USA
ISSN No. 1079-9796

*Published by the Blood Cells Foundation
Established by Springer-Verlag, Inc. in 1975*

groups have generally concluded that homozygosity at the HFE locus is observed in 3-10 individuals per thousand surveyed (2). The classic description of this disease manifesting as "bronze diabetes" partly reflects the pathophysiological targets of iron accumulation (3). Unfortunately, hypogonadotropic hypogonadism, cardiac conduction disorders and chondrocalcinosis also may precede cirrhosis, perhaps the most severe complication of this disease (4). Because of the late age of onset, environmental and sex influence on disease expression, and clinically heterogeneous presentation, hemochromatosis is often diagnosed after significant iron deposition has occurred. Unfortunately for many patients, therapeutic phlebotomy is often initiated after the liver parenchyma has been irreversibly damaged (5).

Over the past twenty years, a variety of studies have examined genetic aspects of hemochromatosis. The initial observation of an HLA-association (6) has been followed by multiple reports of linkage studies which have confirmed the localization of the gene to the short arm of chromosome 6 (7-11). The strong founder effect exhibited by this disease has allowed linkage disequilibrium mapping techniques to be applied in an attempt to accurately position the HFE gene. However, issues related to marker informativeness (12) and density (13) have encumbered the unambiguous localization of the HFE gene to a region readily assailable by conventional gene detection technologies.

A study of the Brittany population, which utilized MHC-derived RFLPs of modest informativeness and HLA-encoded serological determinants as genetic markers, firmly established marker I.82, situated 100 kb proximal to the *HLA-A* locus, as the centromeric limit for the location of the HFE gene (12). However, the enormity of the gene search problem was concomitantly underscored by Jazwinska et al. (14) who demonstrated the strong association between HFE and an allele of the D6S105 microsatellite system, a locus now

known to reside at least 3 megabases (Mb) from the MHC proper (15). The distance between these loci suggests that an extended range of linkage disequilibrium between HFE and chromosome 6 markers exists. A more recent study (13) has used a new statistical method for linkage disequilibrium calculations (16) to examine the associations of polymorphic microsatellites with the HFE gene in a British population. These authors examined 10 VNDRs (variable number of dinucleotide repeats) dispersed over a ~5 cM region and found one, D6S1260, which yielded a significant association with HFE on founder chromosomes. Because of the large gap which exists between this marker and its telomeric neighbor, D6S299 (Figure 1), a definitive placement of the HFE gene was not achievable. Clearly, more markers in the region surrounding D6S1260 need to be analyzed in order to refine the localization of the hemochromatosis critical region.

In the present report, we describe the results of a linkage disequilibrium mapping study carried out for the first time on a large sample of American hemochromatosis chromosomes. In addition to including previously published genetic markers in our analysis, we have utilized a series of recently described polymorphic di- and tetranucleotide repeat polymorphisms which flank the D6S1260 locus. Using several measures of linkage disequilibrium, we refine the centromeric and telomeric borders for the location of the HFE disease gene and discuss the implications of our results with respect to current positional cloning efforts.

MATERIALS AND METHODS

Subjects

Sixty-one unrelated Caucasian hemochromatotic individuals, who had previously undergone liver biopsies, volunteered to participate in our study. Consent forms, approved by the Institutional Review Board of the Pennsylvania State University College of Medicine (IRB Protocol No. 93-140EP) were

reviewed and signed by all participants. Lymphocytes were isolated from twenty milliliters of peripheral blood through the use of a Ficoll gradient and transformed with EBV. DNA was prepared from these cell lines using a non-organic DNA extraction kit (Oncor, Gaithersburg, MD).

Control subjects used in this study have been described (17); they were derived from unrelated Caucasians enrolled in an ongoing quantitative trait loci mapping study. The inclusion of subjects in our study was dependent upon a negative family history for hemochromatosis.

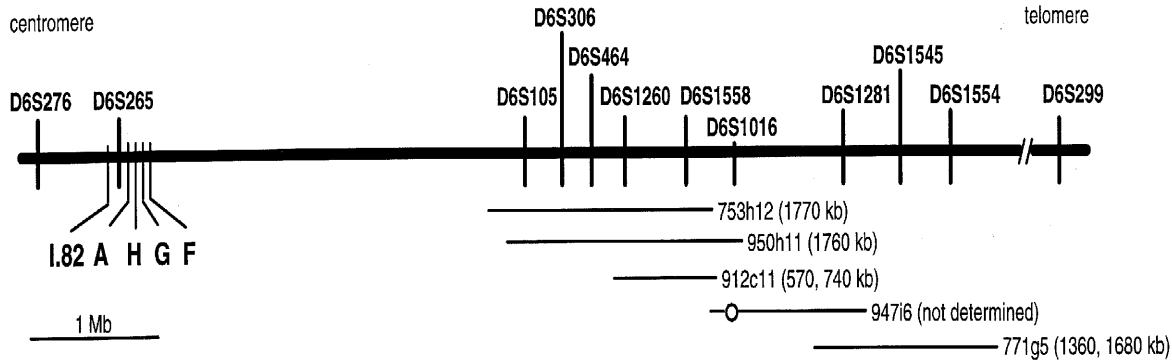


Fig. 1. Scale map of the MHC class I region and distal markers (20,21) used in the study. D6S1260 has been positioned at 700 kb telomeric to D6S105 (13). Both D6S1260 and D6S1558 are contained within YAC 912c11 (containing inserts of 570 and 740 kb); in the figure, D6S1558 has been positioned at the maximum potential distance from the D6S1260 locus. The positions of D6S1016, D6S1281, D6S1545 and D6S1554 are approximations based on insert size and STS content of the YACs. A gap exists in the YAC contig as D6S1554 and D6S299 have not been physically linked. Therefore, D6S299's distance from the *HLA-A* locus is uncertain but is probably greater than 6 Mb. The deletion in YAC 947i6 is indicated by a circle.

Genetic Markers

Polymorphic loci and their associated primer sequences used in this study are presented in Table 1. Some of these markers have appeared in previous studies (13,14) and include D6S105 (18), D6S464 (19), and D6S1260 (13). Certain di- (D6S1558, D6S1545 and D6S1554) and tetranucleotide (D6S1016 and D6S1281) repeat polymorphisms are presented in this study for the first time following their publication (20,21). Marker information can be obtained from the Whitehead Institute-MIT Center for Genome Research World Wide Web server at URL <http://www-genome.wi.mit.edu/>. An updated map of the region between D6S276 and D6S299, including a YAC contig containing the new markers utilized in this report, is presented in Figure 1.

DNA Amplification

Microsatellites were amplified by PCR according to the method of Weber and May (22). Forward primers were labeled with [γ - 32 P]ATP by T4 polynucleotide kinase (New England Biolabs, Beverly, MA). Ten μ l reactions included 20 ng of DNA, 400 nM of each primer, 50 mM KCl, 10 mM Tris, 1.0-1.5 mM MgCl₂, 200 μ mol dNTPs, 1% BSA, and 0.25 U Taq polymerase (Promega, Madison, WI). Reactions were overlaid with mineral oil and amplified using a Perkin-Elmer Cetus 480 Thermocycler. Cycling generally included a "hot-start" denaturation step followed by 27 cycles of denaturation, annealing and elongation with a final 10 minute extension at 72 degrees. Primer annealing temperatures and magnesium concentrations yielding optimal amplification for

each marker pair are presented (Table 1). Products were mixed with a formamide/dye solution, heat denatured and loaded onto a 6% denaturing polyacrylamide gel for analysis. Alleles for each marker were detected after

autoradiography, and band sizes were determined by comparison with an M13 sequencing ladder (U.S. Biochemical, Cleveland, OH).

Table 1. Polymorphic Loci and Associated Marker Information

Locus	Marker Name	GDB ID	Product Size	Type	Tm	MgCl ₂	Primers (5'-3')
D6S276	158.3/158.5	G00-574-089	198-228 bp	(CA) _n	62	1.5 mM	F- TCAATCAAATCATCCCCAGAAG R- GGGTGCAACTTGTTCCTCT
D6S265	AFM101XA1	G00-187-972	122-144 bp	(CA) _n	56	1.5mM	F- ACGTTCGTACCCATTAACCT R- ATCGAGGTAACAGCAGAAA
D6S105	MFD61	G00-178-530	116-138 bp	(CA) _n	55	1.5 mM	F- GAAGGAGAATTGTAATCCG R- GCCCTATAAAATCCTAATTAAC
D6S306	AFM248XH1	G00-188-534	236-250 bp	(CA) _n	56	1.0 mM	F- TGAGAGTTTCAGTGAGCC R- TTTACTTCTGTTCCTTAATG
D6S464	AFM323VB5	G00-199-962	202-230 bp	(CA) _n	50	1.5 mM	F- TGCTCCATTGCACTCC R- CTGATCACCTCGATATTTTAC
D6S1260	CS-5	G00-454-966	144-166 bp	(CA) _n	64	1.0 mM	F- ACTGCTCCTGGCATGGTTG R- GTACATGCCTTGTAAACATC
D6S1558	AFMA192WG9	NA*	251-271 bp	(CA) _n	59	1.5 mM	F- GCTACTTGGGAGGCTGGAC R- CTGGCAGGAGGGCTAGTG
D6S1016	GGAA10G12	G00-695-180	236-268 bp	tetra	56	1.5 mM	F- GCTTAAAATTTAAAAGTGAGTTCC R- CCTGTCAGCTAGAGAGGCAG
D6S1281	GATA89B07	G00-685-863	176-208 bp	tetra	58	1.5 mM	F- GATGCCACGTTTTAAAATGC R- AGAAGCAGCTGTGCTTTGTT
D6S1545	AFMA116ZE1	NA*	225-233 bp	(CA) _n	55	1.5 mM	F- AATCTATGCTCCTGGGTTG R- GAAGTCTGGAATACAGCCTC
D6S1554	AFMA183WB5	NA*	162-176bp	(CA) _n	56	1.5 mM	F- CAACAATAGAAACAGTCCCTTGATG R- CCAGACAGAAATATAAGGCAATTAC
D6S299	AFM217XC7	G00-188-400	210-234 bp	(CA) _n	54	1.0 mM	F- AGGTCATTGTGCCAGG R- TGTCTATGTATACTCCTGAATGCT

* Primer sequences are deposited in the Whitehead Institute Center for Genome Research Databases.

Data Analysis

Genetic markers were manually scored in an independent fashion by at least two individuals. Allele frequencies were determined for the set of disease (n=122) and control (n=88) chromosomes. Statistical methods used to measure the extent of linkage disequilibrium between marker sets have been described. Briefly, P_{excess} was calculated according to the method of Lehesjoki et al. (23). The likelihood-based disequilibrium analysis program, DISLAMB (16), which was kindly provided by Dr. Joseph Terwilliger (University of Oxford, Oxford, UK), was used to calculate lambda (λ)

for each marker.

RESULTS

Genotype Data

The distribution of marker alleles is presented in Figure 2. The sizes of alleles comprising each marker have been translated into simple numerical values with "1" designated as the smallest allele segregating within our population. Throughout this report, locus designation followed by a number denotes a locus-specific allele (e.g., D6S265-1). Table 2 presents the most common alleles found

within the patient collection (disease alleles) and their frequencies within the control group; statistical significance was evaluated by Chi-square analysis. Based on these data, values for P_{excess} , λ (and associated p value) and

homozygosity have been calculated for each marker and are also included in Table 2. Values derived from these three calculations are presented in graphical form in Figure 3.

Allele #	D6S276 ^{a*}		D6S265 ^a		D6S105 ^a		D6S306 ^a		D6S464 ^a		D6S1260 ^a	
	H	C	H	C	H	C	H	C	H	C	H	C
1	15	8	47	10	1	0	0	3	0	0	0	0
2	4	9	1	1	1	0	1	1	1	2	2	2
3	0	0	25	30	1	2	100	52	99	47	9	22
4	15	7	11	2	1	1	9	18	5	4	93	34
5	1	0	21	22	66	25	0	2	2	0	9	11
6	1	2	15	12	6	5	8	10	2	4	2	0
7	8	6	1	1	21	30	1	1	1	0	0	1
8	2	1	0	0	12	15	3	1	6	15	3	4
9	6	5	0	0	9	7			1	1	1	2
10	50	29	1	1	1	0			0	0	2	0
11	17	8	0	0	2	3			0	0	1	5
12	0	0	0	1	1	0			0	0	0	1
13	3	3							3	2		
14									0	2		
15									2	1		
Total	122	78	122	80	122	88	122	88	122	78	122	82

Allele #	D6S1558 ^a		D6S1016 ^b		D6S1281 ^b		D6S1545 ^a		D6S1554 ^a		D6S299 ^{a#}	
	H	C	H	C	H	C	H	C	H	C	H	C
1	0	1	2	0	1	2	17	10	1	0	25	14
2	0	0	3	2	0	1	63	40	0	0	3	3
3	12	8	66	39	1	3	8	3	0	0	4	0
4	1	4	7	4	11	17	25	21	1	1	33	15
5	96	53	6	4	57	27	9	8	68	51	0	0
6	6	14	28	28	37	31			24	12	17	8
7	0	2	9	5	13	4			25	21	32	20
8	2	0	1	3	2	1			3	1	7	7
9	3	4	0	1							0	1
10	0	0									1	0
11	2	0										
12												
13												
14												
15												
Total	122	86	122	86	122	86	122	82	122	86	122	68

Legend
a: size increases by 2 bp per allele
b: size increases by 4 bp per allele
*: allele 1 = 198 bp, allele 2 = 206 bp,
following alleles increase by 2 bp each
#: allele 9 = 226 bp, allele 10 = 234 bp

Fig. 2. Distribution of alleles on hemochromatotic (H) and control (C) chromosomes.

Table 2. HFE Disease Allele Data

Locus	Frequency HFE	Frequency Controls	2x2 Chi-Square p value	P_{excess}	λ (0.05)	p value	Homozygosity (%)
D6S276-10	0.40	0.37	5.9×10^{-1}	0.04	0.00	0.5	20
D6S265-1	0.39	0.13	1.0×10^{-4}	0.30	0.23	1.7×10^{-3}	10
D6S105-5	0.54	0.28	2.0×10^{-4}	0.36	0.27	2.4×10^{-3}	33
D6S306-3	0.82	0.59	3.0×10^{-4}	0.56	0.49	8.6×10^{-4}	67
D6S464-3	0.81	0.60	1.2×10^{-3}	0.53	0.40	7.0×10^{-3}	66
D6S1260-4	0.76	0.42	5.1×10^{-7}	0.59	0.50	1.2×10^{-5}	64
D6S1558-5	0.79	0.62	7.2×10^{-3}	0.45	0.30	2.9×10^{-2}	64
D6S1016-3	0.54	0.45	2.1×10^{-1}	0.16	0.00	0.5	30
D6S1281-5	0.47	0.31	2.7×10^{-2}	0.22	0.13	1.7×10^{-1}	21
D6S1545-2	0.52	0.49	6.9×10^{-1}	0.06	0.00	0.5	26
D6S1554-5	0.56	0.59	6.1×10^{-1}	0.00	0.00	0.5	23
D6S299-4	0.27	0.22	4.5×10^{-1}	0.06	0.00	0.5	11

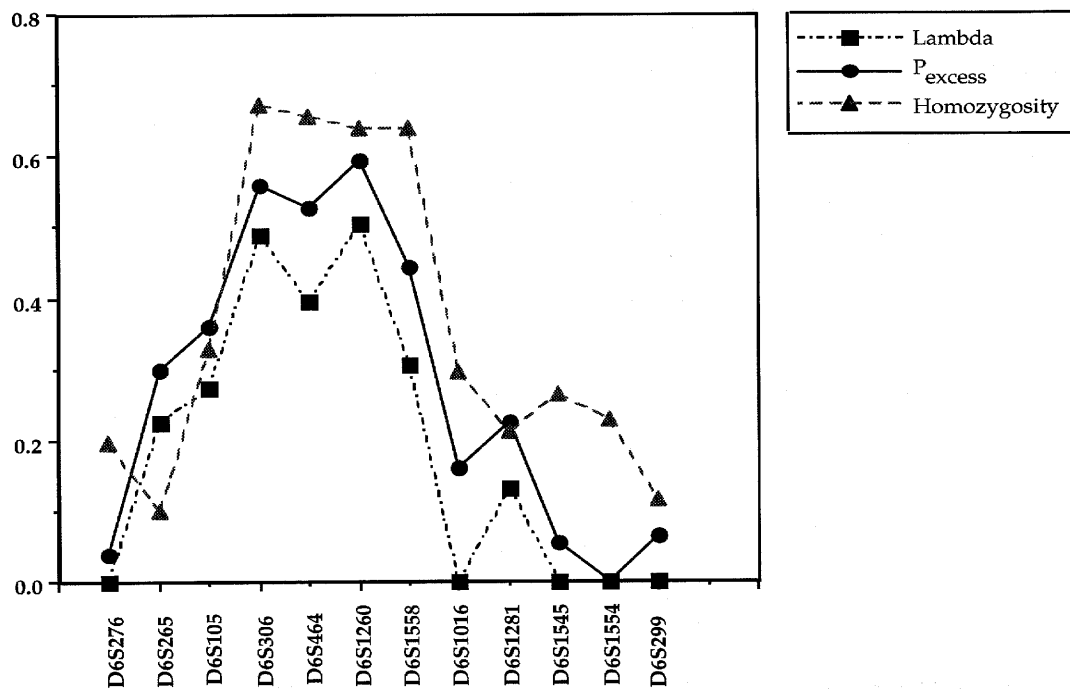


Fig. 3. Values for P_{excess} , lambda, and frequency of homozygosity.

Disease Allele Frequency and the Demonstration of Linkage Disequilibrium

The D6S105 locus, which had been considered a significant landmark for the presence of the HFE locus, yielded an allele 5 (124 bp) frequency of 54% on patient chromosomes [control value of 28% (Table 2)]. This frequency compares well with the frequency of the same allele in both British [66%; control frequency of 26%; (24)] and Australian patient populations [62%; control frequency of 12%; (14)], although a slight diminution was observed. These latter populations elicit a founder effect reflective of a hemochromatosis-specific, ancestral haplotype comprised in part of D6S105-5 and the *HLA-A*-linked VNDR allele, D6S265-1 (25,26). D6S265-1 is very highly correlated with *HLA-A*0301* (27) and thus serves as an economical substitute for *HLA-A* serotyping. Once again, we found a patient frequency of D6S265-1 of 39% (controls: 13%) which compares favorably with the frequencies of 45% and 42% within both the British (26) and Australian (14) hemochromatotic populations, respectively. Based on these two markers, we conclude that the American patient collection comprises a similar proportion of chromosomes carrying the same founder mutation as that found within the previously studied Celtic populations.

Allelic frequencies associated with HFE-carrying chromosomes increased telomeric from the D6S105 locus. Allele frequencies plateaued at approximately 80% from D6S306 through D6S1558, which represents a physical expanse of 1.2 Mb or less. Following D6S1558, a decline in allele frequencies was present through to the D6S299 locus which, due to the observation of a recombination (24), previously represented the telomeric demarcation of the linkage disequilibrium peak.

Homozygosity

The percentage of patients demonstrating homozygosity at each locus is listed in Table 2

and graphically displayed in Figure 3. Homozygosity for D6S265-1 was 10% and for D6S105-5 was 33%. Homozygosity in the disease group remained relatively uniform (averaging 65%) for the D6S306-D6S464-D6S1260-D6S1558 markers and then tapered to 30% for D6S1016 and 11% for D6S299. The frequency of homozygosity of disease alleles on control chromosomes was low for D6S265-1 (3%) and D6S105-5 (2%) and rose to approximately 38% for markers D6S306-3, D6S464-3 and D6S1558-5 (data not shown). In contrast, D6S1260-4 demonstrated only a 20% homozygosity frequency within the control group. The relatively high degree of homozygosity at the D6S306, D6S464 and D6S1558 loci within the controls may reflect the generally poor degree of heterozygosity of each of these three markers evinced through studies on extended French pedigrees (CEPH database site: <http://www.cephb.fr/ceph-genethon-map.html>). To further examine the pattern of allele distribution in our sample, we aligned the hemochromatotic and control patient groups (data not shown) around the markers which exhibited the highest homozygosity in an attempt to determine whether a core genotype was present. The common genotype of D6S464-3/D6S1260-4/D6S1558-5 was observed in 39 patients on 67 chromosomes (54% of the total HFE chromosomes); of this subset, 28 patients (45%) were homozygous at all three loci. Of the other 11 patients, 11 chromosomes carried the core haplotype and all but one of the remaining chromosomes could be converted to this haplotype by allowing for a solitary slippage event. Thus, by patterns of homozygosity alone, a common genotype appears to be associated with HFE in this population.

P_{excess} and the Measure of Linkage Disequilibrium

To further assess the strength of the apparent homozygosity associations, and to compensate for the disease allele frequency within the

control chromosome group, a simple but powerful value, termed P_{excess} , was computed (Table 2 and Figure 3). The P_{excess} statistic (23), which is a measure of linkage disequilibrium, is derived by subtracting the disease allele's frequency within the control group (P_{normal}) from the frequency of the disease allele on disease-bearing chromosomes (P_{affected}) and dividing this value by $[1-(P_{\text{normal}})]$. In populations with a strong founder effect, the theoretical P_{excess} maximum for a marker allele exclusively linked to an unmapped disease locus is 1.0. Heterogeneity will tend to reduce the maximum P_{excess} obtainable. P_{excess} for D6S265 and D6S105 was 0.30 and 0.36, respectively. Moving distally from D6S105, P_{excess} rose with a maximum value of 0.59 for D6S1260. The value of P_{excess} for D6S1558 dropped to 0.45 and continued to fall to 0.16 for D6S1016. The values remained insignificant through D6S299.

λ Likelihood

Lambda (λ) likelihood calculations were performed for each locus (Table 2 and Figure 3) using the program DISLAMB (16), which, in addition to calculating λ , also independently computes chi-square and p values for each marker. The λ statistic is defined as the proportion of increase of a disease-associated allele relative to its population frequency and has been equated to a robust version of the population attributable risk (16,28). Like P_{excess} , λ can achieve a maximum value of 1.0 and has been demonstrated to represent a stringent method for analyzing linkage disequilibrium. Calculating λ requires an estimate of the frequency of a given disease; therefore, a range of values (0.01, 0.05, and 0.067), encompassing generally accepted estimates of the disease gene frequency, were used in the examination of each marker association. We observed little difference in the λ values obtained throughout the above range of frequencies and present our data formulated at 0.05, a frequency expected in both the general populations of the United States (2) and United Kingdom (29). D6S1260 yielded

the highest value for λ of 0.50. Moreover, this was the only marker evaluated in this study to achieve statistical significance ($p \sim 0.00001$) at all estimates of disease allele frequency. Thus, the seemingly high homozygosity rates at flanking loci, when re-examined by λ likelihood statistics, did not achieve statistical significance indicative of a biological association.

DISCUSSION

Successful positional cloning endeavors depend on the precise localization of a disease gene within a well characterized genomic region. Several avenues are available for researchers to pursue these goals. Chromosomal perturbations, such as translocations and deletions, can provide a convenient mechanism to both localize and identify genes of biological significance. When such reagents are not available, linkage studies can guide gene searches but usually can only approximate chromosomal positioning to within 1 to 2 cM. Allelic associations, present when a founder effect can be demonstrated in a population afflicted with a genetic disorder, allow linkage disequilibrium mapping to be implemented in positional cloning efforts. Such methods have been utilized to precisely locate the gene for diastrophic dysplasia (30) and promise to refine the genetic mapping of a variety of inherited disorders, including idiopathic hemo-chromatosis.

Genetic analysis, as it has been applied to localizing the HFE gene, has exhibited a limited yield to date. The highly significant association with the HLA locus, combined with the early reports of "recombinant" individuals, initially supported an MHC class I regional location for the HFE gene. Even a well executed linkage disequilibrium study concurred with a subtelomeric class I positioning for the hemochromatosis locus (12). *In toto*, this information inappropriately focused a variety of investigators to search the gene-dense MHC class I region for HFE candidate genes (31-33). Subsequently, the telomeric positioning of the

HFE locus around the marker D6S105 has been followed by a report which has described attempts to finely position this disease gene by linkage disequilibrium mapping (13).

Within the past several years, many research groups have begun to utilize statistical measures of gametic association, such as P_{excess} , to fine-map disease loci relative to informative chromosomal markers. The best example of the power of these methods was the successful isolation of the gene for an osteochondrodysplasia. The diastrophic dysplasia locus was mapped to within an interval of approximately 70 kb (34) before a genomic sequence was isolated which was subsequently found to encode a novel sulfate transporter (30). In order to apply similar methods to genetic disorders which afflict more heterogeneous population groups, a common haplotype must be present to allow identification of the ancestral mutant chromosome and a dense set of polymorphic markers flanking the region of interest must exist.

Evidence exists to support the notion that recombination rates at the telomeric end of the MHC are non-uniform (35) and it has been suggested that this process may depend on structural differences contained within this subregion (36). We hypothesize that one explanation for the extensive linkage disequilibrium and extended HLA associations is that relatively small populations may have suppressed recombination between the HFE locus and the *HLA* subregion due a proportionately high frequency of *HLA-A* subregional expansions (*HLA-A19*-associated) and contractions (*HLA-A9*-associated) (36). Such "indel" polymorphisms may represent perturbations capable of affecting the initiation of recombinant intermediates during meiosis. In a population such as the American group reported here, with an indeterminate degree of HLA polymorphism and linked structural variability, recombination suppression may have

occurred proportionately less often. This may have allowed for increased exchanges between the HFE locus and flanking markers, thus reducing the genomic expanse of the ancestral genotype in the American population.

Alternative hypotheses exist. Linkage disequilibrium mapping has theoretical concerns (37) which center on unmodelled population genetic phenomena such as drift, mutation, atypical breeding systems and selection, all of which can mask recombination and distort the correlation between physical distance and gametic associations. Considering the proximity of the HLA complex to the hemochromatosis locus and the established role of MHC-encoded determinants in human fecundity (38) and immunity (39), the potential pitfalls of linkage disequilibrium analysis should be fully considered. Nevertheless, the stark discrepancy between genetic (14) and physical mapping data (15) in the MHC to D6S105 region supports our hypothesis.

A continuation of the above statistical approaches may not contribute significantly to further narrowing the HFE gene's residency. However, the continued analysis of new markers flanking the D6S1260 locus on chromosomes harboring ancestral HFE haplotypes may unearth historical recombinational sites which could reduce the genomic expanse demanding analysis. Such an approach, termed recombinant mapping, has been successfully applied to the localization of disease susceptibility genes within the MHC (40). As opposed to genetic analysis, our group is now adopting a comprehensive gene search within the D6S464-D6S1558 interval; a similar approach has already led to the initial characterization of a highly promising HFE candidate gene contained within this region (41). Moreover, a consortium venture may now be required to cogently and efficiently characterize emerging coding sequences with respect to their involvement in genetic iron overload.

ACKNOWLEDGMENTS

This work was supported by grants from the Hackett Foundation (M.J.C.), the March of Dimes (Basic Research Award) (M.J.C.), the NIH [NIDDKD #1 R01 DK43734-01A2 (M.J.C.) and a Penn State College of Medicine Clinical Research Center Grant (#95-43) (G. S. G.). The authors would like to thank Dr. Matrgrit Krikker (President of the Hemochromatosis Foundation) and Ms. Roberta Crawford (President of the Iron Overload Diseases Association, Inc.) for their continued interest and support. This work was performed toward completion of an MD./Ph.D. combined degree (C.P.V.)

REFERENCES

- Halliday JW, Powell LW. Hemochromatosis and other diseases associated with iron overload. In Lauffer RB, ed. *Iron and Human Disease*. Boca Raton: CRC Press, pp. 131-160, 1992.
- Edwards CQ, Kushner JP. Screening for hemochromatosis. *N Engl J Med* 328:1616-1620, 1993.
- Cotran R, Kumar V, Robbins S, eds. *Robbins Pathologic Basis of Disease* (5th edition). Philadelphia: WB Saunders Company, 1994.
- Bothwell TH, Charlton RW, Motulsky AG. Hemochromatosis. In Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The Metabolic and Molecular Bases of Inherited Disease*. New York: McGraw-Hill Publishing Company, pp. 2237-2270, 1995.
- Niederau C, Fisher R, Sonnenberg A, Stremmel W, Trampish HJ, Strohmeyer G. Survival and causes of death in cirrhotic and in noncirrhotic patients with primary hemochromatosis. *N Engl J Med* 313:1256-1262, 1985.
- Simon M, Bourel R, Fauchet R, Genetet B. Association of HLA-A3 and HLA-B14 antigens with idiopathic hemochromatosis. *Gut* 17:332-334, 1976.
- Cartwright GE, Skolnick M, Amos DB, Edwards CQ, Kravitz K, Johnson A. Inheritance of hemochromatosis: linkage to HLA. *Trans Assoc Am Phys* 91:273-281, 1978.
- Simon M, Le Mignon L, Fauchet R, Yaouanq J, David V, Edan G, Bourel M. A study of 609 haplotypes marking for the hemochromatosis gene: (1) mapping of the gene near the HLA-A locus and characters required to define a heterozygous state and (2) hypothesis concerning the underlying cause of hemo-chromatosis HLA association. *Am J Hum Genet* 41:89-105, 1987.
- Lipinski M, Hors J, Saleun JP, et al. Idiopathic hemochromatosis: Linkage with HLA. *Tissue Antigens* 11:471-474, 1978.
- Gasparini P, Borgato L, Piperno A, et al. Linkage analysis of 6p21 polymorphic markers and the hereditary hemochromatosis: localization of the gene centromeric to HLA-F. *Hum. Mol. Genet* 2:571-576, 1993.
- Gruen JR, Goei VL, Summers KM, et al. Physical and genetic mapping of the telomeric major histocompatibility complex region in man and relevance to the primary hemochromatosis gene (HFE). *Genomics* 14:232-240, 1992.
- Yaouanq J, Perichon M, Chorney MJ, et al. Anonymous marker loci within 400 kb of HLA-A generate haplotypes in linkage disequilibrium with the hemochromatosis gene (HFE). *Am J Hum Genet* 54:252-263, 1994.
- Raha-Chowdhury R, Bowen DJ, Stone C, et al. New polymorphic microsatellite markers place the haemochromatosis gene telomeric to D6S105. *Hum Mol Genet* 4:1869-1874, 1995.
- Jazwinska EC, Lee SC, Webb SI, Halliday JW, Powell LW. Localization of the hemo-chromatosis gene close to D6S105. *Am J Hum Genet* 53:347-352, 1993.
- Volz A, Boyle JM, Cann HM, Cottingham RW, Orr HT, Zeigler A. Report of the second international workshop on human chromosome 6. *Genomics* 21:464-472, 1994.
- Terwilliger JD. A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am J Hum Genet* 56:777-787, 1995.
- Plomin R, McClearn GE, Smith DL, et al. DNA markers associated with high versus low IQ: The IQ quantitative trait loci (QTL) Project. *Behavior Genetics*, 24:107-118, 1994.
- Weber JL, Kwitek AE, May PE, Zoghbi HY. Dinucleotide repeat polymorphism at the D6S105 locus. *Nucl Acids Res* 19:968, 1991.
- Stone C, Pointon JJ, Jazwinska EC, et al. Isolation of CA dinucleotide repeats close to D6S105; linkage disequilibrium with haemo-chromatosis. *Hum Mol Genet* 3:2043-2046, 1994.
- Hudson, TJ, Stein LD, Gerety SS, et al. An STS-based map of the human genome. *Science* 270:1945-1954, 1995.
- Dib C, Faure S, Fizames C, et al. The Final Version of the Genethon Human Genetic Linkage Map. *Nature*, in press.
- Weber JL, May PE. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 44:388-396, 1989.
- Lehesjoki AE, Koskiniemi M, Norio R, et al. Localization of the EPM1 gene for progressive myoclonus epilepsy on chromosome 21: linkage disequilibrium allows high resolution mapping. *Hum*

- Mol Genet* 2: 1229-1234, 1993.
24. Raha-Chowdhury R, Bowen DJ, Burnett AK, Worwood M. Allelic associations and homo-zygosity at loci from HLA-B to D6S299 in genetic hemochromatosis. *J Med Genet* 32:446-452, 1995.
 25. Jazwinska, EC, Pyper, WR, Burt, MJ, et al. Haplotype analysis in Australian hemochromatosis patients: Evidence for a predominant ancestral haplotype exclusively associated with hemochromatosis. *Am J Hum Genet* 56:428-433, 1995.
 26. Worwood M, Raha-Chowdhury R, Dorak MT, Darke C, Bowen DJ, Burnett, AK. Alleles at D6S265 and D6S105 define a haemochromatosis -specific genotype. *Brit J Haemat* 86:863-866, 1994.
 27. Worwood M, Raha-Chowdhury R, Darke C. Distribution of alleles at D6S105 and D6S265 with possible HLA haplotype associations. *Tissue Antigens* 44:322-325, 1994.
 28. Delvin B, Risch N. A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* 29:311-322, 1995.
 29. Worwood, M. Genetics of haemochromatosis. *Baillier's ClinHematol* 7:903-918, 1994.
 30. Hästbacka J, de la Chapelle A, Mahtani MM, et al. The diastrophic dysplasia gene encodes a novel sulfate transporter: Positional cloning by fine-structure linkage disequilibrium mapping. *Cell* 78:1073-1087, 1994.
 31. El Kahloun A, Chauvel B, Mauvieux V, et al. Localization of seven new genes around the HLA-A locus. *Hum Mol Genet* 2:55-60, 1993.
 32. Goei VL, Parimoo S, Capossela A, Chu TW, Gruen, JR. Isolation of novel non-HLA gene fragments from the hemochromatosis region (6p21.3) by cDNA hybridization selection. *Am J Hum Genet* 54:244-251, 1994.
 33. Venditti CP, Harris JM, Geraghty DE, Chorney MJ. Mapping and characterization of non-HLA multigene assemblages in the human MHC class I region. *Genomics* 22:257-266, 1994.
 34. Hästbacka J, de la Chapelle A, Kaitila I, Sistonen P, Weaver A, Lander ES. Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. *Nat Genet* 2:204-211, 1992.
 35. Koller, BH, Geraghty, DE, DeMars R, Duvick L, Rich SS, Orr H. Chromosomal organization of the human major histocompatibility complex class I gene family. *J Exp Med* 169:469-480, 1989.
 36. Venditti CP, Chorney MJ. Class I gene contraction within the *HLA-A* subregion of the human MHC. *Genomics* 14:1003-1009, 1992.
 37. Hill WG, Weir BS. Maximum-likelihood estimation of gene location by linkage disequilibrium. *Am J Hum Genet* 54:705-714, 1994.
 38. Ober, C, Elias S, Kostyu DD, Hauck WW. Decreased fecundity in Hutterite couples sharing HLA-DR. *Am J Hum Genet* 50:6-14, 1992.
 39. Tomlinson, IPM, Bodmer WF. The HLA system and the analysis of multifactorial genetic disease. *Trends Genet* 11:463-524, 1995.
 40. Degli-Esposti MA, Andreas A, Christiansen FT, Schalke B, Albert E, Dawkins RL. An approach to the localization of the susceptibility genes for generalized myasthenia gravis by mapping recombinant ancestral haplotypes. *Immuno-genetics* 35:355-364, 1992.
 41. Beutler E, Gelbart T, West C, Kuhl W, Lee P. A strategy for cloning the hereditary hemo-chromatosis gene. *Blood Cells Mol Dis* 21:207-216, 1995.