Am J Hum Genet 37:700–718, 1985

Genetic Analysis of Idiopathic Hemochromatosis Using Both Qualitative (Disease Status) and Quantitative (Serum Iron) Information

J. M. LALOUEL,[1,2] L. LE MIGNON,[3] M. SIMON,[3] R. FAUCHET,[4] M. BOUREL,[5] D. C. RAO,[2] AND N. E. MORTON[6]

SUMMARY

An ongoing family study of idiopathic hemochromatosis in Brittany, France, allowed us to investigate the segregation of this trait and its linkage and association to the *HLA-A* locus in 147 pedigrees, comprising 1,408 individuals with over 900 characterized for relevant biological parameters and typed for HLA. The joint consideration of affection status and serum iron concentration reveals no dominance effect on the latter trait and documents the increased information afforded by the consideration of a biological correlate of liability to affection for disease exhibiting incomplete penetrance. Our overall results are in general agreement with published results on a Utah family study.

INTRODUCTION

Idiopathic hemochromatosis (IH) is a familial disorder classically defined by the triad of cirrhosis of the liver, skin hyperpigmentation, and diabetes mellitus [1]. It is an iron-storage disorder where progressive accumulation of iron in parenchymal cells may lead to these classical symptoms as well as other clinical signs such as hypogonadism, arthropathy, and myocardiopathy. With a positive iron balance of about 2 mg a day [2, 3], decades may pass before an

Received September 11, 1984; revised January 17, 1985.

This work was supported by grants from INSERM (CRL INSERM CNAMTS 1981 to J. M. Lalouel, CRL INSERM-CNAMTS 1982 to M. Simon) and from Fondation Langlois.

[1] Laboratoire d'Anthropologie Biologique, Universite Paris 7, France.

[2] Division of Biostatistics, Washington University School of Medicine, St. Louis, Mo.

[3] Clinique Médicale B, Hôpital Sud, 16, boulevard de Bulgarie, 33022 Rennes Cedex France. Request reprints from M. S.

[4] Centre Régional de Transfusion Sanguine, Rennes, France.

[5] Clinique Médicale A, Centre Hospitalier Universitaire de Rennes, France.

[6] Population Genetics Laboratory, University of Hawaii, Honolulu, Hawaii.

excessive iron storage of about 20–40 g leads to full-blown clinical manifestation; as a consequence, all degrees of severity may be encountered. A number of biological variables are usually found altered: serum iron, transferrin saturation, serum ferritin, iron hepatic concentration, and chelatable iron are generally increased. Indeed, recognition of biological perturbations may lead to early diagnosis before the occurrence of patent clinical symptoms.

A number of family studies have attempted to unravel the mode of inheritance of IH. In fact, every conceivable model has found its partisans, as reviewed in [4]. The identification of association of IH to *HLA-A3* by Simon et al. [5] has clarified this issue, revealing very tight linkage to HLA and providing new support for the recessive hypothesis [6, 7]. With the added consideration of biological variables such as serum iron or transferrin saturation, the question has arisen of a possible expression of the gene among heterozygotes: significant elevations of serum iron, transferrin saturation, or ferritin among unaffected first-degree relatives of affected individuals have been reported (see [4] for review). Moreover, Kravitz et al. [8], by segregation analysis of transferrin saturation in pedigree data, concluded in favor of partial manifestation among heterozygotes.

An ongoing family study is under way in Brittany, France, since 1966. We report here the results of a genetic analysis of 147 kindreds typed for HLA since 1975, with special regard to IH and serum iron, for which an appropriate methodology has been developed. The data, together with some preliminary analysis, are reported in [9].

THE FAMILY STUDY

Protocol of the Survey

Probands for the family study were ascertained from consecutive patients referred to the Cliniques Médicales A and B, Centre Hospitalier Régional, Rennes. Informed consent was sought for screening of first-degree relatives and spouses of probands; whenever feasible, the investigation was extended to more remote relatives.

A control sample of independent individuals was collected through hospitalized patients without hepatopathy or inflammation, hospital staff members, and non-blood relatives of probands from familial data, as in the latter case the correlations between spouses regarding all biological parameters studied were not significant.

Clinical and Laboratory Procedures

Whenever a suspicion of hemochromatosis was entertained for a given patient, on the basis of either noted clinical or biological signs, a thorough screening was undertaken, including: (1) determination of serum iron, total iron-binding capacity, and serum ferritin; (2) whenever feasible, the chelatable iron pool investigated by a Desferrioxamine test; (3) confirmation of the diagnosis sought by performing a liver biopsy for histological examination and determination of iron hepatic concentration; and (4) in instances where the biopsy was not performed, diagnosis established by tolerance of a therapy of weekly phlebotomies over a sufficient time period.

After informed consent, individuals meeting diagnostic criteria were entered as probands in the family study. The same clinical and laboratory screening was performed among the spouse and the first-degree relatives of the probands, although the chelation test and liver biopsy were conducted only when such confirmation was required for diagnosis. HLA typing of *A* and *B* loci was performed on all individuals participating in

the study by the microlymphocytotoxicity method [10]. At a later stage of the study, other markers of chromosome 6 (*DR, Bf, Glo1*) were also determined.

Serum iron and total iron-binding capacity were assayed by the method of Ramsay [11, 12]. The same method was applied in the Desferrioxamine test. Hepatic iron concentration was determined following Barry and Sherlock [13]. These assays were performed at the laboratory of the Clinique Médicale A, Rennes. Ferritin was assayed by radioimmunoassay [14] by C. Beaumont at INSERM U49, Rennes.

MODELS FOR SEGREGATION AND LINKAGE ANALYSES

Analysis of a Dichotomy of Affection Status

Idiopathic hemochromatosis develops slowly as storage iron progressively increases over time. It follows that almost all possible degrees of severity could, in theory, be defined. In such instances, the necessity of disease management often leads clinicians to put forward strict diagnostic criteria, imposing a dichotomy on a continuous diathesis, and, thus, defining normal and affected individuals. As a consequence of the phenotype definition and the possible influence of other factors on the disease process, of which some could be familial, onset and disease penetrance may be function of age as well as familial factors. Were a single gene involved in determining disease susceptibility, such phenotypes would not, in general, exhibit classical Mendelian ratios.

Statistical inference on the mode of inheritance of such discrete phenotypes must then be based on models that attempt to account for such particular aspects of the disease process, as does the mixed model of Morton and MacLean [15] and its subsequent extensions and implementation into the computer program POINTER [16, 17]. Accordingly, we assume that liability to IH is an unobservable variable, x, which results from the joint effects of genotype at an autosomal locus, g, a multifactorial transmissible component, c, and random, nontransmitted environmental factors, e, such that $x = g + c + e$, with variance $V = G + C + E$. Affection occurs whenever x exceeds some threshold, which may be made a function of age and sex by assigning specific thresholds to various liability classes, defined from corresponding morbid risks appropriate for the reference population. The mean, u, and variance, V, of x are arbitrary, taken as 0 and 1, respectively. This model involves the following parameters for the major locus: the gene frequency, q, of the allele, a, leading to increased disease liability; the distance between homozygote means on the liability scale, or displacement, t; and the position of the heterozygote mean relative to the two homozygote means, or dominance d; in addition, one may consider general transmission probabilities [18], τ_1, τ_2, and τ_3, which specify, respectively, the probabilities of the three genotypes AA, Aa, and aa transmitting the normal allele, A. The multifactorial transmissible component requires two additional parameters, H and z, where $H = C/V$ is the heritability in childhood while z is the ratio of adulthood and childhood heritabilities. Complete specification of the model can be found in [15–17].

This model allows tests of hypotheses regarding: (1) the extent of familial aggregation; (2) the existence of a major gene, its dominance relation, and agreement of transmission with Mendelian expectations; (3) residual family

resemblance; and (4) heterogeneity among mating types. Such tests are carried out by the likelihood ratio criterion, where twice the difference between the log-likelihoods of an hypothesis H_1 and a subhypothesis H_0 involving k independent parameter restrictions is asymptotically distributed as a χ^2 with k degrees of freedom when H_0 is true.

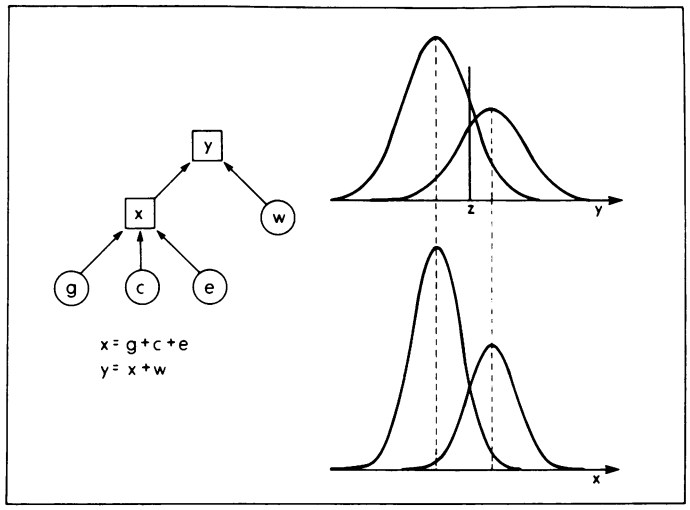
The Adjunct Consideration of a Biological Indicator

There is some arbitrariness in the previous disease definition that leads to a serious loss of information in inferring both the mode of inheritance and genetic risks. Under the major gene hypothesis, penetrance is low in younger individuals, and individuals classified as normal do not contribute much information. More information would be expected of a phenotype definition that could take into account various grades of affection, and some authors have proposed the definition of major and minor iron load in IH (e.g., [8]). This approach, however, would again require the specification of rather arbitrary diagnostic criteria and could lead to some confusion of both severity and duration of the disease process.

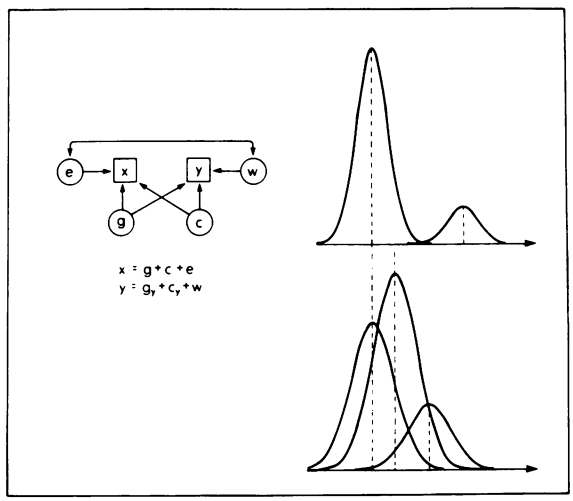
The underlying disease process being continuous rather than discrete, a natural alternative, for an investigation into etiology, would consist in defining a continuous variable that would be a good indicator of intrinsic liability to disease rather than a measure reflecting the actual course, duration, or severity of the disease; that is, we seek a variable, or a combination of variables, which would be primarily altered in genotypes at risk prior to affection, rather than a variable whose alteration would be subsequent to affection per se.

Morton and MacLean [15] proposed the following model for such a situation (fig. 1*A*). Let x, as defined earlier, be an observable biological correlate of liability to affection, the latter being now defined by y, and assume that the relationship between x and y is such that $y = x + w$, with $\text{var}(y) = \text{var}(x) + \text{var}(w)$. Affection is defined by a threshold on the y-scale, and w represents random environmental effects, the correlation between x and y being equal to $\{\text{var}(x)/[\text{var}(x) + \text{var}(w)]\}^{1/2}$. A consequence of this formulation is that, as $\text{var}(w) > 0$, we have $\text{var}(y|g) > \text{var}(x|g)$, which constrains the overlap of liability distributions within major genotypes to be greater than that of the corresponding distributions for the biological indicator. This constraint did not appear appropriate in the present instance, as will be shown further, which led us to propose an alternate formulation (fig. 1*B*).

Let x denote a biological indicator and y denote the liability of affection such that disease occurs whenever y exceeds some threshold, and $x = g + c + e$ with expectation $E(x) = u$ and variance $V = G + C + E$ as before, while $y = g_y + \alpha c + \beta e + w$ with zero expectation and unit variance V_y such that $V_y = G_y + \alpha^2 C + \beta^2 E + W$. While the parameters defined earlier refer implicitly to x, we are led to define new parameters for y: dominance, d_y; displacement, t_y; heritability, $H_y = \alpha^2 C/V_y$; and environmental covariance, $E_y = \beta E$. This model assumes that a major locus, a multifactorial transmissible component, and nontransmitted environmental factors exert specific effects on x and y without the previous unnecessary constraint. It preserves the general probability for-



DISCRETE - CONTINUOUS BIVARIATE PHENOTYPE: MODEL I



DISCRETE - CONTINUOUS BIVARIATE PHENOTYPE: MODEL II

FIG. 1.—*A* (*top*), Two genetic models for a bivariate discrete/continuous phenotype. *B* (*bottom*), Two genetic models for a bivariate discrete/continuous phenotype.

mulation previously used [16, 17], particularly regarding the two-dimensional integration. Different displacements and dominance parameters can be considered on x and y. The latter generalization also seems necessary in view of the fact that one can conceivably envision the case of a recessive condition for which a biological trait leads to recognition of heterozygotes, at least in probability.

This model allows segregation analysis on a bivariate phenotype consisting of a dichotomy of affection status and/or a quantitative trait. The conditional phenotype probabilities required are given in the APPENDIX.

Linkage Analysis

Both association to *HLA-A3* and linkage to the *HLA-A* locus have been reported ([5, 7, 8, 19]; see [4] for a review). Therefore, linkage analysis has been performed with the computer program LINKAS [20], which accommodates association in terms of coupling frequencies, that is, conditional probabilities of the IH allele given alleles at the test locus. To keep computations manageable, the model used for linkage analysis does not make provision for multifactorial inheritance. However, it was extended to allow proper handling of a bivariate phenotype as defined earlier.

GENETIC ANALYSIS: RESULTS

Segregation Analysis of Affection Status

Morbid risks for IH vary with age and sex, the disease being more frequent in males than in females. This was taken into account in the analysis by estimating morbid risks for age- and sex-specific liability classes. Le Mignon [9] reported age at first diagnosis for 220 male and 43 female probands, of whom only three were then under age 20. Omitting nonprobands under that age, we divided each sex into four age groups (20–39, 40–59, 60–79, 80+), and assigned the frequencies of the 1975 French national census. Alexandre [21] estimated the prevalence in the general population as 16/14,000 = .0011. This led to the following estimated cumulative risks at the midpoint of the ith age class [16]: .00098, .00357, .00541, and .00559 in males and .00019, .00064, .00104, and .00109 in females.

No generally valid solution has yet been found to account for ascertainment in the analysis of extended familial structures. For the present analysis, we have followed the pointer logic presented in [16], where a pedigree is resolved into its component nuclear families and the probability density of each sibship is written conditional on the parental phenotypes, the probability of ascertainment through the sibs (if appropriate), and the phenotypes of possible individuals or pointers outside the nuclear family who may have led to ascertainment. An example is given in figure 2. Thus, 147 pedigrees yielded 297 nuclear families comprising 1,408 individuals, including 745 offspring, 922 distinct individuals with affection status known, and 915 individuals with serum iron concentration determined.

This approach requires the estimation of a probability of ascertainment under incomplete selection. The ascertainment probability, π, among families in the sampling frame was estimated as follows. The study was conducted over 15 years of a nominal 70-year generation. There were 151 probands in the family study among 274 total probands. Therefore, $\pi = (15/70)/(151/274) = .118$. This is in reasonable agreement with the estimate of $.061 \pm .041$ from the distribution of probands among affected in 42 sibships with one or more probands. Since the latter estimate has such a large standard error, we used $\pi = .118$.

The segregation analysis of affection status was rather inconclusive. There was patent familial aggregation, as well as evidence for a major effect, but Mendelian transmission was rejected, and no resolution could be obtained between dominant and recessive hypotheses. Under the mixed model, all max-

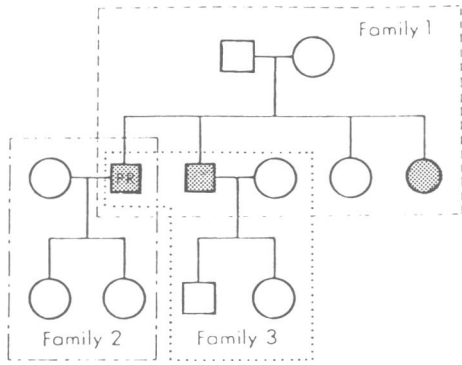


FIG. 2.—Resolution of a pedigree into component nuclear families while taking into account ascertainment. *Family 1*: incomplete selection; *family 2*: complete selection; *family 3*: complete selection with proband (*PR*) pointing to his brother. Affection is indicated by *shading*.

imum likelihood solutions were obtained for a residual variance of liability $E = 0$. The analysis under a generalized single-locus model without residual family resemblance, while favoring a codominant Mendelian model ($d = 0.35$), led to rejection of Mendelian transmission. These results indicate that either the model or the assumed prevalences are inappropriate to account for these data.

The Choice of a Biological Indicator

As mentioned above, a biological indicator of liability to affection should, for the purpose of a formal genetic analysis under the models entertained presently, be independent of both the actual duration and the severity of the disease; otherwise, it could lead to misclassification, in probability, of genotypes at risk and thereby distort estimates of genetic parameters, in particular segregation ratios.

With several quantitative traits available for the present analysis, we had to choose between the alternatives of either constructing a discriminant function with regard to a classification of affection status or selecting a particular trait for segregation analysis. In view of the above discussion and the assumptions that a discriminant analysis would imply concerning dominance, the latter approach was adopted for this preliminary investigation involving a new model.

The measures of both the chelatable iron pool and the hepatic iron concentration, available only on some definitely affected individuals, are indeed indicators of disease rather than of liability. Serum ferritin is closely correlated to the size of the body iron stores [22, 23].

Transferrin saturation is the ratio of two variables: serum iron concentration, divided by the total iron-binding capacity of transferrin (TIBC). The latter is assayed by progressive addition of iron to a sample until the binding capacity of transferrin is fully saturated; it is therefore an indirect measurement of serum transferrin concentration that is not independent of serum iron concentration, and such ratios are known to have undesirable statistical properties. An arcsine transformation may stabilize its variance, but there remains the additional

problem that more than 50% of the probands in the present study exhibit 100% saturation. The serum transferrin concentration is inversely related to the size of the body iron stores [24]; moreover, transferrin synthesis occurs primarily in the liver and is therefore dependent on hepatic function and, possibly, also reduced with increasing serum ferritin concentrations [25, 26]. It is probably the best biological discriminant of affection status in IH, as its increase may result from both increased serum iron and decreased serum transferrin associated with liver damage.

In contrast, serum iron concentration is reportedly elevated in IH for minimally increased iron stores [27], and the serum iron concentrations in males and females are similar despite differences in the average body iron stores [28]. Without ignoring the clinical significance of other parameters, we have therefore chosen serum iron as a biological indicator of liability for further genetic analyses.

The Distribution of Serum Iron in Controls and Probands

Characteristics of the distribution of serum iron in controls and probands are reported in table 1. There is no significant difference in means between the two sexes within both controls and probands, although they are slightly lower in females than in males, as often reported [28]. A larger variance is found in male probands than in controls. Positive skewness, at the limit of significance in male controls, is quite significant among male probands. Kurtosis, not significant in controls, is significantly increased in male probands.

Polynomial regression of serum iron on age revealed a significant quadratic relation in male controls, leading to the equation: (expected) serum iron = $19.515 + 0.204$ age $- 0.00350$ age^2; serum iron increases moderately with age through adolescence (from 21.0 μmol/l at age 5 to 22.3 at age 20), remains fairly stable until age 40, and then decreases progressively with advancing age, to about 19.0 by age 60. The same general trend is also found in male probands. This trend is rather weak when referred to the variance of serum iron, as attested by a multiple correlation coefficient of .19. Family and control data were, however, adjusted for age according to this equation and standardized by the estimated residual variance for further analysis. In contrast, no significant association was found between serum iron and age in females, whether controls or probands.

After standardization, male and female controls were pooled and revealed persistent skewness ($\beta_1 = 0.186$, $P < .05$). Although such marginal skewness is unlikely to simulate segregation of a major gene, evidence for some degree of expression in heterozygotes could be sensitive to distributional assumptions. Therefore, an analysis of distributions allowing for a mixture of two underlying distributions as well as a power transformation to remove skewness, following [29], was carried on the control sample (table 2). Skewness is significant when one underlying distribution is assumed ($\chi^2_1 = 5.07$); it is not significant under the assumption of two underlying distributions ($\chi^2_1 = 1.77$). Evidence for a mixture of distributions with allowance for a power transformation is marginally significant ($\chi^2_1 = 5.70$), the proportion of individuals in the higher

TABLE 1

HEMOCHROMATOSIS: MOMENTS OF BIOLOGICAL VARIABLES IN CONTROLS AND PROBANDS

Sample	Sex		Age (yrs)	Serum iron (μM/l)	TIBC* (μM/l)	Saturation (%)	LIBC† (μM/l)	Ferritin (μM/l)	Desferal (μM/l)
Controls	Male	Mean	33.3	21.4	63.3	30.7	46.6	178.5	24.0
		σ	16.0	6.2	8.0	11.7	10.8	83.1	5.3
		No.	166	166	86	86	86	94	31
	Female	Mean	39.4	20.5	66.6	31.3	46.3	79.6	23.6
		σ	15.6	6.8	7.7	11.9	10.9	89.8	11.7
		No.	235	233	136	137	137	190	20
Probands	Male	Mean	43.8	44.8	50.5	89.7	5.7	2465.8	161.7
		σ	9.6	8.2	7.9	13.3	8.1	2748.3	82.8
		No.	230	220	217	218	218	78	176
	Female	Mean	45.6	42.3	48.4	88.8	5.4	1953.2	127.0
		σ	12.4	6.2	7.1	10.4	5.4	3278.7	99.2
		No.	44	43	43	43	43	21	36

* TIBC = total iron-binding capacity of transferrin.
† LIBC = latent iron-binding capacity of transferrin.

TABLE 2

SERUM IRON IN CONTROLS: ANALYSIS OF SKEWNESS AND COMMINGLING

No. distributions	Power transformation	−2 ln L	V	u	t	q	p
1	No	952.34	0.988	−0.000	(0)	(0)	(1)
1	Yes	947.27	0.989	−0.048	(0)	(0)	0.421
2	No	943.34	1.295	0.554	1.721	0.339	(1)
2	Yes	941.57	1.440	0.696	1.982	0.335	1.500

NOTE: −2 ln L: minus twice the natural logarithm of the likelihood. V: total variance; u: overall mean; t: distance between means of two distributions; q: admixture parameter; p: parameter of power transform. Parameters values between parentheses are held constant.

distribution being estimated as 0.11, around a mean of 1.6 σ. Previous genetic analyses [8, 19] that favored the hypothesis of some degree of expression of the IH gene in heterozygotes, also led to an estimated gene frequency of .05. The present results being compatible with this hypothesis, the segregation analysis was first performed without, then with, a power transformation.

Segregation Analysis of IH and Serum Iron

The first step of the analysis revealed that a model relating x and y through the sole consideration of an additional random component w ([15]; fig. 1*A*) was not appropriate for the present data (table 3). For the general model, it led to the rejection of the hypothesis of Mendelian transmission ($\chi^2_1 = 48.83$), and clearly a common displacement on both scales could not simultaneously account for appropriate disease penetrances and the proper degree of overlap of serum iron distributions in the two homozygous classes. Thus, for the best-fitting Mendelian model, the lifetime morbid risks (over age 80) in normal homozygotes, heterozygotes, and homozygotes for the IH gene are: .003, .013, and .456 in males, .0004, .0019, and .223 in females; the corresponding risks of affection are: .603, .207, and .189 in males and .363, .162, and .474 in females. All available data point to a higher penetrance in IH homozygotes in males and a much lower proportion of sporadic cases.

This led us to entertain the alternate model discussed earlier (fig. 1*B*), used in all subsequent analyses. Because serum iron exhibited some degree of skewness that, we have noted, could be compatible with heterozygous expression, the segregation analysis was first performed without skewness-removing transformation in order to avoid a possible reduction of evidence for partial manifestation. Essential results are reported in table 4.

The best-fitting Mendelian model favors recessivity for both serum iron and liability to affection: dominance parameters, when iterated, do not differ significantly from zero ($\chi^2_2 = 1.93$). More remarkable is the fact that, by contrast with the previous analysis, the Mendelian hypothesis is quite acceptable ($\chi^2_1 = 0.00$). Mendelian transmission was therefore assumed for subsequent tests. Taking the recessive hypothesis with Mendelian transmission as general model (row 3), intermediate ($d = d_y = 0.5$) and dominant ($d = d_y = 1.0$)

TABLE 3

HEMOCHROMATOSIS: SEGREGATING ANALYSIS OF AFFECTION STATUS AND SERUM IRON UNDER MODEL 1A

Model	Transmission	$-2 \ln L + C$	V	u	d	t	q	H	W	τ_2
Mixed model	General	2732.06	1.675	−1.112	0.225	4.427	0.0362	0.448	1.059	0.261
Mixed model	Mendelian	2780.39	1.692	−0.949	0.171	4.190	0.0482	0.550	1.164	(0.5)
Single locus	General	2763.27	1.449	−0.959	0.327	4.637	0.0320	(0)	1.097	0.225
Single locus	Mendelian	2832.94	1.278	−0.451	0.263	4.139	0.0490	(0)	1.240	(0.5)
Multifactorial	.....	2924.56	3.519	−2.146	(0)	(0)	(0)	0.684	1.279	...

NOTE: Parameters are defined in the text; values between parentheses are held constant.

TABLE 4

HEMOCHROMATOSIS: SEGREGATION ANALYSIS OF AFFECTION STATUS AND SERUM IRON UNDER MODEL 1B

Model	Transmission	Dominance	$-2 \ln L + C$	V	u	d	t	q	H	d_s	t_s	τ_2
Mixed model	General ...	Estimated	2510.29	1.575	0.590	−0.090	3.196	0.0597	0.184	−0.170	7.0	0.49
Mixed model	Mendelian .	Estimated	2510.29	1.574	0.590	−0.090	3.196	0.0597	0.184	−0.184	7.0	(0.5)
		$d = d_s = 0$	2512.22	1.566	0.379	(0)	3.438	0.0597	0.185	(0)	7.0	(0.5)
		$d = 0, d_s = 0.5$	2512.87	1.558	0.372	(0)	3.443	0.0605	0.186	(0.5)	10.0	(0.5)
		$d = d_s = 0.5$	2794.47	1.507	0.176	(0.5)	6.452	0.0020	0.060	(0.5)	12.5	(0.5)
		$d = d_s = 1$	2798.58	1.575	0.202	(1.0)	3.239	0.0020	0.076	(1.0)	6.5	(0.5)
Single locus	General ...	Estimated	2522.00	1.563	0.610	−0.105	3.153	0.0597	(0)	−0.100	7.0	(0.5)
		$d = d_s = 0$	2523.98	1.568	0.375	(0)	3.423	0.0597	(0)	(0)	7.5	(0.5)
		$d = d_s = 0.5$	2795.77	1.520	0.170	(0.5)	6.408	0.0020	(0)	(0.5)	12.6	(0.5)
		$d = d_s = 1.0$	2800.22	1.586	0.193	(1.0)	3.218	0.0020	(0)	(1.0)	7.0	(0.5)
Multifactorial	.....	...	3168.32	1.534	−0.618	(0)	(0)	(0)	0.722	(0)	(0)	...

models are overwhelmingly rejected (twice the logarithm of the odds are 282.25 and 286.36, respectively). The best-fitting models, whether Mendelian or general, did not show evidence of either environmental correlation between serum iron and disease liability or residual familial correlation for disease liability. This is not altogether surprising in view of the limited overlap between liability distributions within major genotypes. Indeed, Reich et al. [30] showed by computer simulations that no resolution of residual family resemblance could be expected in such instances. By contrast, there is support in favor of some residual familial correlation for serum iron ($H = 0.18$, $\chi^2_1 = 11.76$). Still, it must be noted that conclusions are not altered when this factor is ignored. As for the major gene hypothesis, its support is also overwhelming ($\chi^2_3 = 656.10$).

It appeared, however, that the likelihood surface was very flat with respect to dominance on the liability scale, d_y, as reflected in row 4 of table 4. Therefore, the adjunction of serum iron to the phenotype definition, under the present model, and the prevalence assumed here, do not help in resolving this issue. Given that the inferred IH gene appears recessive for serum iron levels, it seems more likely that it would also be so for liability to disease, but this argument stems more from biological considerations than from statistical support.

Let us examine the parameter estimates for the best-fitting Mendelian model. Displacements, relative to the within-genotype standard deviations of x and y, are 3.04 and 7.0, respectively; this explains why the previous model could not apply to the present data. The estimated gene frequency, .06, is in good agreement with the frequency reported by Cartwright et al. [19] for a set of Mormon pedigrees.

The available evidence does not support a hypothesis of partial expression of the IH gene for serum iron. It is therefore unlikely that skewness of serum iron in controls results from some degree of heterozygous expression. As serum iron as such exhibits greater variance among male probands, the same analysis was repeated after a power transformation to remove skewness and stabilize the variance [15]. Results were essentially unaltered by this transformation, leading to the following estimates for the best-fitting Mendelian model: $V = 1.19$, $u = 0.31$, $d = 0.0$, $t = 2.91$, $q = .06$, $H = .19$, $d_y = 0.0$, $t_y = 7.0$, $H_y = 0$, and $E_y = 0$. Segregation at the major locus accounts for 3% and 15% of the total variance in x and y, respectively, which leads to a correlation of .07 between serum iron and liability. This shows that, although the major gene accounts for the joint segregation of serum iron and affection, it does not imply an important correlation between serum iron and disease liability. The previous model, by its emphasis on correlation between x and y, could not account for the observed results.

Linkage and Association with HLA

Significant associations have been reported between hemochromatosis and the HLA antigens A3, B7, and B14 [4]. Association appears to be primarily with the antigen A3, with a relative risk of 6.61 for these data [9], the association with B7 and B14 being in accordance with their well-documented linkage

disequilibrium with A3 [9]. Le Mignon et al. [31] also reported evidence of a significant association with A11, which could have been unnoticed previously because of the strong association already exhibited with the more prevalent allele *A3*. In view of the known cross-reactions between A3 and A11, this observation deserves some attention. Previous linkage analyses under the assumption of linkage equilibrium have yielded maximum lod scores for a recombination value of $\theta = 0$ [8, 19].

Analysis of linkage between hemochromatosis and the *HLA-A* locus was performed on 147 pedigrees including 930 HLA-typed individuals, with the computer program LINKAS [20]. This program allows both linkage and association. Rather than estimating sex-specific recombination rates, we have assumed a ratio of female-to-male genetic distances of 2, as seems reasonable in the HLA region [32]. The model therefore includes as parameters the male recombination rate, θ_m, and coupling frequencies c_i, $i = 1, \ldots, 6$, where c_i is the conditional probability of the IH gene given the ith allele at the *HLA-A* locus. Noniterated coupling frequencies are constrained to be equal. It was appropriately extended to accommodate the present model, although multifactorial transmission was not taken into account. The parameters of the model were taken as reported above after power transformation, with H neglected. LINKAS allows consideration of a maximum of six alleles at the test locus. Hence, *HLA-A3* and *-A11* were specifically coded, while other alleles were randomly recoded following Ott [33].

The results of this analysis are reported in table 5. The evidence for linkage is very strong at a maximum likelihood estimate of $\theta_m = 0.011 \pm 0.0078$, with a likelihood-ratio $\chi^2 = 174.77$, which corresponds to a lod score of 37.95. The association with *A3* is very significant ($\chi^2_1 = 148.28$). The association with *A11* also appears significant ($\chi^2_1 = 10.80$), adding support to previous claims [31]. Similar results are obtained when affection status alone is considered, with a maximum lod score of 34.59 at $\theta_m = 0.0073$.

As the hypothesis $\theta_m = 0$ could be rejected with these data ($\chi^2_1 = 6.14$), we examined lod scores at $\theta_m = 0$ in each pedigree in order to identify possible evidence of recombination. Four pedigrees give lod scores close to 1.0 and

TABLE 5

IDIOPATHIC HEMOCHROMATOSIS: ANALYSIS OF LINKAGE AND ASSOCIATION WITH HLA-A

Hypothesis	−2 ln L	θ_m	c_{A3}	c_{A11}	c_0
θ_m, $\hat{c}_{A3}$, $\hat{c}_{A11}$	8447.68	0.011 ±0.0078	0.168 ±0.0051	0.080 ±0.0057	(0.036) ...
θ_m, $\hat{c}_{A3}$	8458.48	0.012	0.168	(0.039)	(0.039)
θ_m, $\hat{c}_{A11}$	8595.96	0.013	(0.059)	0.081	(0.059)
$\theta_m = 0$, $\hat{c}_{A3}$, $\hat{c}_{A11}$	8453.82	(0)	0.167	0.081	0.036
$\theta_m = 0.5$, $\hat{c}_{A3}$, $\hat{c}_{A11}$	8622.20	(0.5)	0.172	0.088	0.035

NOTE: Iterated parameters are indicated by a "hat." Constrained values are reported within parentheses. When not iterated, coupling frequencies are constrained to be equal and such that $\Sigma_i c_i p_i = q$ over all alleles at the test locus, where c_i and p_i are the coupling frequency and gene frequency of the ith test locus allele, and q is the gene frequency of the hemochromatosis gene; c_0 denotes all the other alleles.

share the following features: the only informative sibship contains two affected individuals that share one *HLA* haplotype, the other haplotypes being identical at the *HLA-A*, but differing at the *HLA-B* locus; these latter haplotypes have been coded as distinct haplotypes, which will tend to overestimate recombination between *HLA-A* and IH whenever a recombinational event may have occurred between the *HLA-A* and *-B* loci. Had such families been coded without the *B* locus, they would have been uninformative about recombination and θ_m would no longer be significantly different from zero. Another pedigree yielded a lod score of 1.64; most of the evidence for recombination comes from two individuals who are concordant for *HLA-A* but discordant for disease status. Two unaffected individuals in this pedigree exhibit elevation of both serum iron and ferritin levels while chelatable iron is moderately elevated in one of them. This pedigree deserves careful reexamination before concluding this as a possible recombinational event.

DISCUSSION

The mode of inheritance of idiopathic hemochromatosis has long been a source of controversy. The present analysis, when restricted to the segregation analysis of a classification of affection status, was indeed inconclusive. It illustrates difficulties associated with the very definition and analysis of progressive disorders whose population prevalence and age- and sex-specific morbid risks, in the absence of a thorough population-based, epidemiological investigation, are at best only very approximately estimated. Analysis under the mixed model with age- and sex-specific thresholds led to inflated estimates of heritability and subsequent failure to distinguish, with the available methodology, between dominant and recessive inheritance. Careful examination of the data revealed it can, in part, be attributed to an underestimate of morbid risks in females as compared to males. While the sex ratio is 4.2:1 among probands, it is 2.4:1 in secondary cases. Resorting to simpler models, such as the generalized single-locus model, is no palliative: here, it favored a Mendelian model with partial expression in heterozygotes ($d = 0.35$), but transmission frequencies departed from Mendelian expectations.

In the absence of more reliable population data, and given the relative arbitrariness of a boundary to separate affected from normal individuals, there was little motivation to refine estimates of morbid risks and repeat the analysis of a disease classification. Rather, our analysis focused on a biological correlate of liability to disease and linkage to the *HLA-A* locus.

The greater information content of quantitative traits as opposed to disease classification implying incomplete penetrance, for segregation analysis, has been emphasized and investigated [34]. When several biological indicators are available, as was the case here, a choice must be made among several possible strategies: combining these indicators into a single function prior to segregation analysis [35], analyzing a multivariate phenotype [36], or carefully selecting a particular variable by considering the physiopathology of the disease at hand. We opted for the last alternative because the first one may blur differences in heterozygous effects of the presumed IH gene on each trait as well as the

primary or secondary nature of their association with the disease process, while the multivariate approach is impractical for complex models and simultaneous consideration of affection status.

Ferritin and chelatable iron reflect duration and severity of disease rather than intrinsic liability to disease. While transferrin saturation, when considered singly, may indeed be the best discriminant between patent cases and normal controls, it is a ratio variable involving in the denominator an indirect estimate of serum transferrin concentration. Transferrin synthesis in the liver is most likely to be affected by the degree of fibrosis of the liver. For these reasons, serum iron appeared to be a better indicator of liability to disease and was therefore selected for further analysis.

In the present situation where ascertainment operates through a definition of affection status, segregation analysis requires combination of both aspects of phenotype definition rather than substitution of a biological variable for affection status. Two alternative formulations of a bivariate, discrete/continuous model were applied to the data. Tests of the transmission parameter τ_2 proved useful in recognizing that the model in figure 1*A* was not appropriate for these data. This led to the development of an alternate model (fig. 1*B*) that proved adequate to account for these observations and that could be of value in other, similar instances where disease definition is complemented by a biological correlate.

The adjunct consideration of a quantitative trait helped to resolve the issue of whether heterozygotes are at an increased risk of affection. The estimated dominance for liability to affection was compatible with a recessive hypothesis for affection status. For serum iron, the data were quite informative about the dominance parameter and led to the conclusion of no partial expression of the IH gene in heterozygotes for serum iron levels. This lends additional support to a recessive hypothesis for affection status. The significant increase of mean serum iron in unaffected relatives of probands compared with controls that some have reported (see [4] for references) may result either from nonidentification of preclinical cases or from the slight but significant residual familial correlation detected in our study. Segregation analysis of transferrin saturation in a large Mormon pedigree [8] could not resolve between two competing hypotheses with dominance parameters of $d = 0.17$ and $d = 0.61$, respectively. Linkage analysis clearly favored the former hypothesis, but joint estimation of recombination and segregation parameters did not modify appreciably estimates of the latter parameters [8].

Linkage analysis confirmed the very tight linkage already reported between IH and the *HLA-A* locus, as well as the association with *A3*. In addition, it confirmed a suspected association with *A11* [31], which may point to yet stronger association of IH with an antigenic determinant common to these two complex antigens with known cross-reactivity. The recombination rate in males was estimated as $.011 \pm .0078$ when using affection status and serum iron as a phenotype and as $.0073 \pm .0072$ when using affection status alone, the corresponding lod scores being 37.95 and 34.59, respectively. This shows that, as opposed to the situation encountered in segregation analysis, most of the

information on tight linkage can be provided by a definition of affection status alone, at least in instances where the estimated proportion of sporadic cases is very low. While the hypothesis of complete linkage ($\theta_m = 0$) could be rejected with these data, there was no clear identification of recombinants. The most suggestive evidence, which came from two individuals concordant for HLA but discordant for the disease phenotype, calls for a careful reexamination of this pedigree.

Our overall conclusions are in good agreement with those of the Utah family study [8, 19]. These two studies, however, were conducted in two populations that may differ in their epidemiological characteristics. Some further differences are the biological variables studied and some methodological aspects already discussed. In the present study, serum iron was selected as a biological indicator of intrinsic liability to disease; it exhibited recessive inheritance. In the Utah family study, the emphasis instead was on transferrin saturation. Presumably because it integrates some of the disease process, it behaved more like an indicator of realized liability. By contributing to the identification of heterozygotes, it may have added some linkage information in extended pedigrees. Further analyses of the biological parameters, including transferrin saturation, singly or in combination with other parameters [37], will be required to complete the genetic and biological analysis of idiopathic hemochromatosis in the present French population.

APPENDIX

CONDITIONAL PHENOTYPE PROBABILITIES UNDER THE NEW BIVARIATE DISCRETE/CONTINUOUS MODEL

Phenotype definition consists of a classification of affection status and a biological correlate of liability to disease. The biological correlate is such that $x = g + c + e$, with variance $V = G + C + E$, with model specifications given in [15]. In particular, we recall that r denotes the parent-offspring correlation between multifactorial transmissible variables, which is also their degree of relationship. Affection status is defined by a threshold on the liability scale, y, which, for convenience of derivation and implementation, we can formulate, in simply notational departures from figure 1*B*, as $y = g_y + \alpha c + \beta e + w$. The major gene effect, g_y, involves specific dominance and displacement parameters d_y and t_y, respectively; α and β are two scalars, and w is an independent random variable $N(0, W)$. Mean and variance of y are arbitrary, taken as 0 and 1, respectively. From these specifications, it follows that $V_y = G_y + \alpha^2 C + \beta^2 E + W = 1$.

Calculation of phenotype probabilities will require expressions of conditional densities of x given major genotype and multifactorial transmissible effects, as already given in [16], and of conditional probabilities of affection status given such effects *and* x, if measured. The latter will be tails of normal distributions whose moments must be obtained for parents, pointers to parents, and offspring or pointers to offspring.

If i denotes the ith major genotype with corresponding effects on x and y denoted g_i and g_{yi}, respectively, the required moments for a parent are

$$E(y|x, i, c) = g_{yi} + \beta(x - g_i) + (\alpha - \beta)c$$

$$\text{var}\,(y|x, i, c) = W$$

if x is known, else

$$E(y|i, c) = g_{yi} + \alpha c$$
$$\text{var } (y|i, c) = \beta^2 E + W \ .$$

For a pointer with major genotype i and degree of relationship r to a parent P, we have, with $A = (1 - r^2)C + E$,

$$E(y|x, i, c_p) = g_{yi} + \{[(1 - r^2)\alpha C + \beta E](x - g_i) + r(\alpha - \beta)Ec_p\}/A$$
$$\text{var } (y|x, i, c_p) = V_y - G_y - \{(1 - r^2)\alpha^2 C^2 + [2\alpha\beta\,(1 - r^2) + r^2\alpha^2]\,CE + \beta^2 E^2\}/A$$

and

$$E(y|i, c_p) = g_{yi} + r\alpha c_p$$
$$\text{var } (y|i, c_p) = (1 - r^2)\,\alpha^2 C + \beta^2 E + W \ .$$

For an offspring, or a pointer to the offspring, with major genotype i, F and M denoting the parents and letting $B = (1 - 2r^2)C + E$, we have

$$E(y|x, i, c_F, c_M) = g_{yi} + \{[(1 - 2r^2)\alpha C + \beta E](x - g_i) + r(\alpha - \beta)E\,(c_F + c_M)\}/B$$
$$\text{var } (y|x, i, c_F, c_M) = V_y - G_y - \{(1 - 2r^2)\alpha^2 C^2 + 2[(1 - 2r^2)\alpha B + r^2\alpha^2]CE + \beta^2 E^2\}/B$$

and

$$E(y|i, c_F, c_M) = g_{yi} + r\alpha(c_F + c_M)$$
$$\text{var } (y|i, c_F, c_M) = (1 - 2r^2)\alpha^2 C + \beta^2 E + W \ .$$

REFERENCES

1. Trousseau A: *Clinique Médicale de l'Hôtel-Dieu de Paris*, 2nd ed. Paris, Ballère, 1865, p 672
2. Finch S C, Finch CA: Idiopathic hemochromatosis, an iron storage disease. *Medicine* 34:381–430, 1955
3. Pollycove M: Hemochromatosis, in *The Metabolic Basis of Inherited Disease*, 4th ed, edited by Stanbury JB, Wyngaarden JB, Fredrickson DS, New York, McGraw-Hill, 1978, pp 1127–1164.
4. Simon M, Alexandre JL, Fauchet R, Bourel M: The genetics of hemochromatosis. *Prog Med Genet* 4:135–168, 1980
5. Simon M, Pawlotsky Y, Bourel M, Fauchet R, Genetet B: Hemochromatose idiopathique: maladie associée à l'antigène tissulaire A3?. *Nouv Presse Med* 4:1432, 1975
6. Simon M, Alexandre JL, Bourel M, Lemarec B, Scordia C: Heredity of idiopathic hemochromatosis: a study of 106 families. *Clin Genet* 11:327–341, 1977
7. Simon M, Bourel M, Genetet B, Fauchet R: Idiopathic hemochromatosis. Demonstration of recessive transmission and early detection by family HLA typing. *N Engl J Med* 297:1017–1021, 1977
8. Kravitz K, Skolnick M, Cannings C, et al.: Genetic linkage between hereditary hemochromatosis and HLA. *Am J Hum Genet* 31:601–619, 1979
9. Le Mignon L: Etude génétique de l'hemochromatose idiopathique: le point à propos de 147 enquêtes familiales comportant un groupage HLA. M.D. thesis, Faculté de Médecine de Rennes, Rennes, France, 1982

10. MITTAL KD, MICKEY MR, SINGAL DP, TERASAKI PI: Serotyping for homotransplantation XVIII. Refinement of microdroplet lymphocyte cytotoxicity test. *Transplantation* 6:913–927, 1968
11. RAMSAY WNM: The determination of iron in blood plasma or serum. *Clin Chim Acta* 2:214–220, 1957
12. RAMSAY WNM: The determination of the total iron-binding capacity of serum. *Clin Chem Acta* 2:221–226, 1957
13. BARRY M, SHERLOCK S: Measurement of liver iron concentration in needle-biopsy specimens. *Lancet* 1:100–103, 1971
14. ADDISON GM, BEAMISH MR, HALES CM, HODGKINS M, JACOBS A, LLEWELLIN P: An immunoradiometric assay for ferritin in the serum of normal subjects and patients with iron deficiency and iron overload. *J Clin Pathol* 25:326–329, 1972
15. MORTON NE, MACLEAN CJ: Analysis of family resemblance. III. Complex segregation of quantitative traits. *Am J Hum Genet* 26:489–503, 1974
16. LALOUEL JM, MORTON NE: Complex segregation analysis with pointers. *Hum Hered* 31:312–321, 1981
17. LALOUEL JM, RAO DC, MORTON NE, ELSTON RC: A unified model for complex segregation analysis. *Am J Hum Genet* 35:816–826, 1983
18. ELSTON RC, STEWART J: A general model for the genetic analysis of pedigree data. *Hum Hered* 21:523–542, 1971
19. CARTWRIGHT GE, EDWARDS CQ, KRAVITZ K, ET AL.: Hereditary hemochromatosis. Phenotypic expression of the disease. *N Engl J Med* 301:175–179, 1979
20. MORTON NE, LALOUEL JM: Resolution of linkage for irregular phenotype systems. *Hum Hered* 31:3–7, 1981
21. ALEXANDRE JL: Etude de l'hérédité dans l'hémochromatose. A propos de 106 enquêtes familiales. M.D. thesis, Faculté de Médecine Rennes, Rennes, France, 1975
22. JACOBS A: Serum ferritin, in *Iron Metabolism and Thalassemia*, edited by BERGSMA D, CERAMI A, PETERSON CM, GRAZIANO JH, New York, Alan R. Liss, 1977, pp 97–104.
23. WALTERS GO, MILLER FM, WORWOOD M: Serum ferritin concentration and iron stores in normal subjects. *J Clin Pathol* 26:770–772, 1973
24. MILDERS MS, COOK JD, FINCH CA: The influence of food iron absorption on the plasma iron level in idiopathic hemochromatosis. *Acta Haematol (Basel)* 60:65–75, 1978
25. MORTON AG, TAVILL AS: The role of iron in the regulation of hepatic transferrin synthesis. *Br J Haematol* 36:383–394, 1977
26. MORTON A, HAMILTON SM, RAMSDEN DB, TAVILL AS: Studies on regulatory factors in transferrin metabolism in man and the experimental rat, in *Plasma Protein Turnover*, edited by BIANCHI R, MARIANI G, Baltimore, University Park Press, 1976, pp 165–177
27. BOTHWELL TH, COHEN I, ABRAHAMS OL, PEROLD SM: A familial study in idiopathic hemochromatosis. *Am J Med* 27:730–738, 1959
28. BOTHWELL TH: *Iron Metabolism in Man*. Oxford, England, Blackwell, 1979
29. MACLEAN CJ, MORTON NE, ELSTON RC, YEE S: Skewness in commingled distributions. *Biometrics* 32:695–699, 1976
30. REICH T, RICE J, CLONNINGER CR: The detection of a major locus in the presence of multifactorial variation, in *Genetic Research Strategies for Psychobiology and Psychiatry*, edited by GERSHON ES, MATTHYSSE S, BREAKEFIELD XO, CIARANELLO RD, Pacific Grove, Calif., Boxwood Press, 1981, pp 353–367.
31. LE MIGNON L, SIMON M, FAUCHET R, ET AL.: An HLA-A11 association with the hemochromatosis allele? *Clin Genet* 24:171–176, 1983
32. KEATS B: Genetic mapping: chomosomes 6-22. *Am J Hum Genet* 34:730–742, 1982
33. OTT J: A simple scheme for the analysis of HLA linkage in pedigrees. *Ann Hum Genet* 42:255–257, 1978
34. MACLEAN CJ, MORTON NE, LEW R: Analysis of family resemblance. IV. Operational characteristics of segregation analysis. *Am J Hum Genet* 27:365–384, 1975

35. NAMBOODIRI KK, ELSTON RC, GLUECK CJ, FALLAT R, BUNCHER CR, TSANG R: Bivariate analysis of cholesterol and triglyceride levels in families in which probands have type IIb lipoprotein phenotype. *Am J Hum Genet* 27:454–471, 1975
36. GOLDIN LR, ELSTON RC, GRAHAM JH, MILLER CH: Genetic analysis of von Willebrand's disease in two large pedigrees: a multivariate approach. *Am J Med Genet* 6:279–293, 1980
37. DADONE MM, KUSHNER JP, EDWARDS CQ, BISHOP DT, SKOLNICK MH: Hereditary hemochromatosis. Analysis of laboratory expression of the disease by genotype of 18 pedigrees. *Am J Clin Pathol* 78:196–207, 1982

Erratum

In the paper "Tourette Syndrome: Clinical and Psychological Aspects of 250 Cases," by D. E. Comings and B. G. Comings (*Am J Hum Genet* 37:435–450, 1985), numbers that appeared in table 3 on p. 441 need correction. They are corrected below.

TABLE 3

TYPE AND FREQUENCY OF MOTOR AND VOCAL TICS THROUGHOUT THE COURSE OF TS

	No.	%		No.	%
Motor tics					
Facial	206	82	Shoulder	79	32
Eyeblinking	140	56	Arms	57	23
Mouth-opening	94	38	Hands	37	15
Facial grimacing	72	29	Flexing-clasping	12	5
Rolling eyes	22	9	Piano fingers	15	6
Squinting eyes	10	4	Copropraxia	7	3
Opening eyes	7	3	Other	15	6
Closing eyes while driving	7	3	Diaphragm	31	12
Sticking tongue out	14	6	Legs and feet	51	21
Licking lips	9	4	Kicking	17	7
Licking shoulder	7	3	Hopping-skipping	11	7
Biting tongue	3	1	Flexing feet	7	3
Looking at the sun	3	1	Other	20	8
Grinding teeth	7	3	Hitting self	17	7
Head and neck	157	63			
Hair out of eyes	31	12			
Horizontal head tic	50	20			
Vertical head tic	23	9			
Chin on shoulder	8	3			
Vocal tics					
Throat-clearing	139	56	Humming	14	6
Grunting	79	32	Yelling-screaming	17	7
Sniffing	37	15	Blow-out breath	16	6
Spitting	22	9	Suck-in breath	11	4
Barking	21	8	Whistling	5	2
Snorting	16	6	Other	52	21
Squeeking	15	6	No vocal tics	9	4
Coughing	17	7			