Am. J. Hum. Genet. 59:1149–1162, 1996

Consanguinity and the Sib-Pair Method: An Approach Using Identity by Descent Between and Within Individuals

Emmanuelle Génin and Françoise Clerget-Darpoux

Unité de recherche d'Epidémiologie Génétique, INSERM U155, and Institut National d'Etudes Démographiques, Paris

Summary

To test for linkage between a trait and a marker, one can consider identical marker alleles in related individuals, for instance, sibs. For recessive diseases, it has been shown that some information may be gained from the identity by descent (IBD) of the two alleles of an affected inbred individual at the marker locus. The aim of this paper is to extend the sib-pair method of linkage analysis to the situation of sib pairs sampled from consanguineous populations. This extension takes maximum advantage of the information provided by both the IBD pattern between sibs and allelic identity within each sib of the pair. This is possible through the use of the condensed identity coefficients. Here, we propose a new test of linkage based on a χ^2. We compare the performance of this test with that of the classical χ^2 test based on the distribution of sib pairs sharing 0, 1, or 2 alleles IBD. For sib pairs from first-cousin matings, the proposed test can better detect the role of a disease-susceptibility (DS) locus. Its power is shown to be greater than that of the classical test, especially for models where the DS allele may be common and incompletely penetrant; that is to say for situations that may be encountered in multifactorial diseases. A study of the impact of inbreeding on the expected proportions of sib pairs sharing 0, 1, or 2 alleles IBD is also performed here. Ignoring inbreeding, when in fact inbreeding exists, increases the rate of type I errors in tests of linkage.

Introduction

A possible way to detect linkage between a trait and a marker consists in showing a positive correlation between the concordance for the trait and the concordance for the marker in sib pairs (Penrose 1935). In other words, this method, referred to as the "sib-pair method," detects linkage when affected sibs share alleles identical by descent (IBD) for the marker more often than expected by chance (Haseman and Elston 1972; Day and Simons 1976; Suarez 1978). Under the hypothesis of independence between trait and marker, affected sibs are expected to have 0, 1, and 2 marker alleles IBD in the respective proportions of 1/4, 1/2, and 1/4. If a departure from these proportions in the proper direction is found, we can suspect linkage. Several tests have been proposed to detect a statistically significant departure (for a review, see Blackwelder and Elston 1985). The method does not require the specification of the mode of inheritance of the disease and can be applied to multifactorial diseases. In particular, it has been used to study linkage with markers in the HLA region. Because this genetic system is highly polymorphic, inferences concerning the IBD relations between the observed alleles among sibs are easy and typing of the parents is generally not necessary. For instance, linkage with HLA was found in leprosy (De Vries et al. 1976) and in insulin-dependent diabetes mellitus (Cudworth and Woodrow 1975). In light of the IBD distribution of haplotypes among affected sibs, it is also possible to make some inferences concerning the mode of inheritance of the disease. Thomson and Bodmer (1977) derived the expected proportions of affected sibs sharing 0, 1, and 2 alleles IBD under a dominant and a recessive biallelic model. These proportions depend on the frequency of the disease-susceptibility (DS) allele and permit discrimination between a recessive and a dominant model, given data on a marker located close to the DS locus. Under some assumptions, the frequency of the DS allele can be estimated.

One major assumption in these studies is that the population from which affected sib pairs are sampled is in Hardy-Weinberg proportions. In particular, it is supposed that there is no inbreeding. The possibility of IBD for the two homologous alleles at a locus within an individual (i.e., the possibility of autozygosity at a locus, is excluded). However, as illustrated by homozygosity mapping (Lander and Botstein 1987), this possibility of autozygosity can be revealing in the study of rare recessive diseases. For this reason, we propose to extend the affected-sib-pair method to assess not only at the IBD pattern between sibs but also within each sib of the pair.

Received January 16, 1996; accepted for publication August 23, 1996.

Address for correspondence and reprints: Dr. Emmanuelle Génin, INSERM U155, Château de Longchamp, Bois de Boulogne, 75016 Paris, France.

© 1996 by The American Society of Human Genetics. All rights reserved.

0002-9297/96/5905-0025$02.00

We use the concept of identity states (Gillois 1964; Jacquard 1972; Karigl 1982). These define the possible IBD states between the four alleles at a locus of two individuals. This concept permits a complete description of the IBD relationships between and within individuals. Identity states should not be confused with the term "identity by state" (Lange 1986), which denotes the fact that two alleles are identical but not necessarily by descent (i.e., it is possible that they are not copies of the same ancestral allele). In this paper, we are concerned only with IBD relationships among alleles between individuals (noted "IB state") and within individuals (noted "IW states"). When considering jointly IBD between and within individuals, we will use the notation "IBW states" instead of "identity states," to avoid confusion.

We first use the concept of IBW states to study the impact of inbreeding on linkage analyses performed by the sib-pair method. We derive the expected IB proportions under the null hypothesis of no linkage and show the departure from the expected proportions of 1/4, 1/2, and 1/4 for sharing, respectively, 0, 1, and 2 alleles IBD when there is inbreeding. In the second part of the paper, we show the gain in the power for linkage detection when jointly taking into account the identity of alleles between and within individuals.

Definition of IBW States

For two individuals I_1 and I_2, ≤ 15 IBW states can be defined to specify the IBD relationships between the four alleles of these individuals at a given locus. Following the method of Thompson (1974), let $\mathbf{a} = (\mathbf{a}_{11}, \mathbf{a}_{12}, \mathbf{a}_{21}, \mathbf{a}_{22})$ be the set of genes of these two individuals, the first subscript indicating the maternal or paternal gene and the second subscript indicating the sibling. The first gene, $\mathbf{a}_{11}$, is labeled 1 and the next three genes are labeled so that two IBD genes are given the same label. If maternal and paternal genes are not distinguishable, then sets (1, 2, 1, 2) and (1, 2, 2, 1) are not distinguishable, and there are nine distinct states (S_i, $1 \leq i \leq 9$) (see table 1). The probabilities of these IBW states (p_i, $1 \leq i \leq 9$) depend on the genealogical relationship between the pair of individuals.

Derivation of the Expected IB-State Probabilities for Two Sibs in a Consanguineous Population: Impact for Tests of Linkage

In the following section, we will consider only pairs of sibs. The probability of IBW states for pairs of sibs depends on the kinship and inbreeding coefficients of their parents. For example, if one considers two sibs whose parents are not related, only three of the nine states have probabilities >0, and these three states correspond to the three IB states: S_9, 0 allele IBD (IB = 0) with probability 1/4; S_8, 1 allele IBD (IB = 1) with probability 1/2; and S_7, 2 alleles IBD (IB = 2) with probability 1/4. If the parents of the sibs are related, some more states are possible and in particular states where both alleles of an individual are IBD (states S_1, S_2, S_3, S_4, S_5, and S_6). Table 1 gives the probability of IBW states for two sibs as a function of the mean inbreeding coefficient α of the population (Jacquard 1970, 1972) from which these sibs are sampled (for derivation, see appendix A). To obtain the probabilities of IB states, IBW-state probabilities have to be pooled, clearly illustrating that information is lost when reducing the nine configurations to three. Two sibs share zero alleles IBD (IB = 0) if they are in the IBW states S_2, S_4, S_6, or S_9 and the probability is $Z = p_2 + p_4 + p_6 + p_9$. Two sibs share one allele (IB = 1) if they are in the IBW states S_3, S_5, or S_8 and the probability is $Y = p_3 + p_5 + p_8$. The probability for two sibs of sharing two alleles IBD (IB = 2) (states S_1 and S_7) is then $X = p_1 + p_7$.

Table 1

Probability of IBW States for a Pair of Sibs Sampled from a Population with a Mean Inbreeding Coefficient α

IBW State	Probability
$S_1(1, 1, 1, 1)$	$(5/8)\alpha^3 + (1/4)\alpha^2 + (1/8)\alpha$
$S_2(1, 1, 2, 2)$	$(1/8)\alpha^4 - (1/4)\alpha^3 + (1/8)\alpha^2$
$S_3(1, 1, 1, 2)$	$-(1/4)\alpha^3 + (1/4)\alpha$
$S_4(1, 1, 2, 3)$	$-(1/8)\alpha^4 + (3/8)\alpha^3 - (3/8)\alpha^2 + (1/8)\alpha$
$S_5(1, 2, 2, 2)$	$-(1/4)\alpha^3 + (1/4)\alpha$
$S_6(1, 2, 3, 3)$	$-(1/8)\alpha^4 + (3/8)\alpha^3 - (3/8)\alpha^2 + (1/8)\alpha$
$S_7(1, 2, 1, 2)$	$(1/8)\alpha^4 - (7/8)\alpha^3 + (1/8)\alpha^2 + (3/8)\alpha + (1/4)$
$S_8(1, 2, 1, 3)$	$-(1/4)\alpha^4 + (5/4)\alpha^3 - (5/4)\alpha^2 - (1/4)\alpha + (1/2)$
$S_9(1, 2, 3, 4)$	$(1/4)\alpha^4 - \alpha^3 + (3/2)\alpha^2 - \alpha + (1/4)$

When comparing the IB state proportions X, Y, and Z with the ones expected without inbreeding, 1/4, 1/2, and 1/4, we note that inbreeding leads to a decrease in the proportion of pairs with zero alleles IBD and to an increase of pairs with two alleles IBD (see table 2). If the mean inbreeding coefficient α is <0.1, the proportion of sib pairs with one IBD allele stays approximately equal to 0.5.

We studied the impact of ignoring inbreeding, when it exists, on linkage tests based on sib pairs. We considered the three statistical tests of linkage studied by Blackwelder and Elston (1985) and derived their statistic under the hypothesis $H_{0\alpha}$ of no linkage and presence of inbreeding.

The t_1 test is based on the proportion of sib pairs with two marker alleles IBD (IB = 2). This test consists in comparing the observed proportion (X) of sib pairs among the N sib pairs of the sample with IB = 2 with the expected proportion ($E[X] = 1/4$) under the null hypothesis H_0 of no linkage and no inbreeding. The statistic $t_1 = (X - E(X))/\sqrt{3/16N}$ is tested as a standard normal deviate.

Table 2

Expected Number of Sib Pairs in the Three IB Categories at the Marker Locus in a Sample of *N* Sib Pairs

IB Category	Expected No. of Pairs, Assuming Inbreeding ($\alpha \geq 0$)	Expected No. of Pairs, Assuming No Inbreeding ($\alpha = 0$)
IB = 0	$(N/4)[(1/2)\alpha^4 - 2\alpha^3 + (7/2)\alpha^2 - 3\alpha + 1]$	$N/4$
IB = 1	$(N/2)[-(1/2)\alpha^4 + (3/2)\alpha^3 - (5/2)\alpha^2 + (1/2)\alpha + 1]$	$N/2$
IB = 2	$(N/4)[(1/2)\alpha^4 - \alpha^3 + (3/2)\alpha^2 + 2\alpha + 1]$	$N/4$

NOTE.—These expectations are given under the hypothesis of no linkage between the marker and a disease-susceptibility locus, assuming that the population from which sib pairs are sampled is consanguineous (first column) or not (second column).

Under $H_{0\alpha}$, t_1 follows a normal distribution with mean

$$\mu = \sqrt{\frac{N}{3}}\left(\frac{1}{2}\alpha^4 - \alpha^3 + \frac{3}{2}\alpha^2 + 2\alpha\right)$$

and variance

$$\sigma^2 = \frac{1}{3}\left(\frac{1}{2}\alpha^4 - \alpha^3 + \frac{3}{2}\alpha^2 + 2\alpha + 1\right) \times \left(-\frac{1}{2}\alpha^4 + \alpha^3 - \frac{3}{2}\alpha^2 - 2\alpha + 3\right).$$

The type I error of the test for a nominal value of 5% may be obtained by computing the probability for t_1 to exceed 1.645 under $H_{0\alpha}$. It is thus the probability for the standard normal distribution to exceed $(1.645 - \mu)/\sigma$.

The mean test t_2 is a test based on the mean number of marker alleles shared IBD by the two sibs. This test compares the observed mean number of alleles shared IBD by the sibs ($m = 2X + Y$) with the expected mean number ($E[m] = 1$) under H_0 (no linkage and no inbreeding). The statistic $t_2 = (m - E(m))/\sqrt{1/2N}$ is tested as a standard normal deviate.

Under $H_{0\alpha}$, t_2 follows a normal distribution with mean

$$\mu = \sqrt{2N}\left(\frac{1}{4}\alpha^3 - \frac{1}{2}\alpha^2 + \frac{5}{4}\alpha\right)$$

and variance

$$\sigma^2 = \left(\frac{1}{2}\alpha^4 - \frac{1}{2}\alpha^3 + \frac{1}{2}\alpha^2 + \frac{9}{2}\alpha + 3\right) - \frac{1}{2}\left(\frac{1}{2}\alpha^3 - \alpha^2 + \frac{5}{2}\alpha + 2\right)^2.$$

As for the t_1 test, the type I error of the t_2 test can be evaluated as a function of α by computing the probability for the standard normal distribution to exceed $(1.645 - \mu)/\sigma$.

The third test is a χ^2 goodness-of-fit statistic (test IB) (with 2 df) that consists in comparing the observed number of sib pairs in each of the three categories (IB = 0, 1, and 2) and the expected number under H_0 ($N/4$, $N/2$, and $N/4$). Under the hypothesis $H_{0\alpha}$, the IB statistic follow a noncentral χ^2 with 2 df and noncentral parameter λ. The noncentral parameter depends on the sample size N and on α:

$$\lambda = N\alpha^2(209 - 250\alpha + 391\alpha^2 - 320\alpha^3 + 179\alpha^4 - 54\alpha^5 + 9\alpha^6).$$

For each of these tests, type I errors have been reported on figure 1 as the function of the mean inbreeding coefficient α of the population, considering sample sizes N of 100 and 30 sib pairs. The type I errors of the three tests increase rapidly with α. The increase is greater for samples of $N = 100$ sib pairs than for samples of

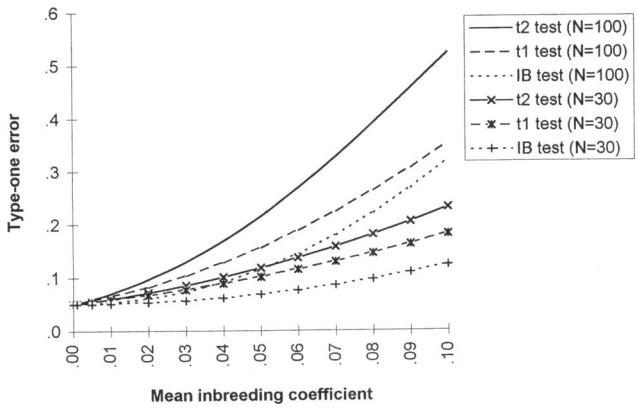

Figure 1 Type I errors of three tests of linkage, based on affected sib pairs as a function of the mean inbreeding coefficient for a nominal value of 5%. The type I error is derived for samples of 100 sib pairs and 30 sib pairs.

$N = 30$ sib pairs. If α is 0.005, type I errors are 5.6%, 5.9%, and 5.1% for the t_1, t_2, and IB tests, respectively, if $N = 100$, and 5.4%, 5.5%, and 5.0%, if $N = 30$. If α is 0.01 and $N = 100$, they are 6.4%, 7.0%, and 5.2%, respectively. For values of $\alpha < 0.01$, type I errors may be considerably increased so that ignoring inbreeding when it exists may lead to a false claim of linkage. The IB test appears to be more robust than the two other tests.

Use of the IBW States When Testing for Linkage on Data of Affected Sib Pairs with Related Parents

Data on Affected Sib Pairs with Related Parents

Above, we considered sib pairs sampled from a consanguineous population with mean inbreeding coefficient α. In this sample, there was a mixture of sibs whose parents were related to different degrees. Here, our aim is to take maximum advantage of inbreeding to detect linkage. For this purpose, we will restrict our attention to the case of sib pairs whose parents have a given relationship, namely, first-cousins. We assumed a sampling scheme focusing on such inbred sib pairs. Illustrations will be given for the case of sib pairs from a first-cousin marriage, but derivations can be performed with the formula, whatever the degree of relationship between the parents of the sibs. The choice of first-cousin marriages was motivated by the fact that, in some countries, in particular Muslim countries, such data may be easily collected. Moreover, it was the sampling scheme considered by Lander and Botstein (1987) for homozygosity mapping. Since inbred sib pairs are expected to be found in populations where mating between relatives is frequent, there may exist in the population a remote consanguinity. We considered that sib pairs from first-cousin matings are sampled from a population with a mean inbreeding coefficient α. First, we studied the power of tests when concerned with such inbred sib pairs. In the last part of the paper, we study the impact of ignoring the remote consanguinity on the tests.

Model

Assume that the presence of an allele D with frequency q increases the susceptibility to a disease. The penetrance of the genotype DD, Dd, and dd are f_1, f_2, and f_3, respectively. Information on a marker locus located at a negligible recombination fraction from the DS locus is assumed to be available. This marker locus is supposed to be highly polymorphic so that two identical alleles are IBD. Considering pairs of individuals with a given genealogical relationship, we can calculate the observed proportions of each IBW state and compare them to the proportions expected under the null hypothesis of no linkage. These expected proportions depend only on the genealogical relationship between the individuals and can be computed using the algorithm of Karigl (1981). However, in this algorithm, the remote consanguinity in the population is ignored. We have extended this algorithm to take it into account (see appendix B). In table 3, the expected IBW-state probabilities are given for the case of a sib pair from a first-cousin marriage sampled from a population with mean inbreeding coefficient α.

Expected IBW-State Probabilities for Pairs of Affected Individuals under a Genetic Model

If we consider a pair of individuals (I_1, I_2) affected by the disease, we can derive the probability p_i' of each IBW state S_i at the marker locus, given that the individuals are affected.

$$p_i = P(S_i \mid I_1 \text{ and } I_2 \text{ affected}) = \frac{P(I_1 \text{ and } I_2 \text{ affected} \mid S_i)p_i}{P(I_1 \text{ and } I_2 \text{ affected})}.$$

Taking into account all the possible genotypes G_1 and G_2 of individuals I_1 and I_2 at the disease locus, we obtain

$$p_i' = \frac{\sum_{G_1,G_2} P(G_1,G_2 \mid S_i)p_i P(I_1 \text{ affected} \mid G_1)P(I_2 \text{ affected} \mid G_2)}{\sum_{G_1,G_2} P(I_1 \text{ affected} \mid G_1)P(I_2 \text{ affected} \mid G_2) \sum_{i=1}^{9} P(G_1,G_2 \mid S_i)p_i}. \quad (1)$$

The probabilities $P(G_1,G_2 \mid S_i)$ are given in table 4. They depend on the frequency q of the DS allele D.

In light of the genealogical relationship between the two members of the pair, it is possible to derive the respective probabilities of each IBW state under different disease models. In table 5, the IBW-state probabilities p_i' are given as a function of the parameters at the DS locus. Note that the probabilities p_i' do not depend on the penetrances (f_1, f_2, f_3) but on the ratio of penetrances ($x = f_2/f_1$ and $s = f_3/f_1$). The probability for the two individuals to share 0, 1, or 2 alleles IBD (probability of IB states) may also be obtained by summing the IBW-state probabilities (see appendix B).

For different frequencies q of the DS allele, Thomson and Bodmer (1977) derived the expected proportion of affected sib pairs sharing 0, 1, and 2 alleles IBD in the case of a recessive or a dominant disease. To obtain these probabilities, in a first step, the possible genotypes of the parents of the sibs were determined, and computations were made conditionally on these genotypes. Here, we can obtain the same results by using the probabilities of the sibs genotypes conditionally on their IBW states. With this conditioning, more general cases can be considered than were discussed by Thomson and Bodmer (1977). The expected proportions of each type of af-

Table 3

Probabilities of IBW States for a Pair of Sibs from a First-Cousin Marriage in a Population with Mean Inbreeding Coefficient α, under the Null Hypothesis of No Linkage

IBM State	Probability	No. of Alleles
$S_1(1, 1, 1, 1)$	$p_1 = (1/4)\alpha^3 + (1/2)\alpha^2 + (15/64)\alpha + (1/64)$	1
$S_2(1, 1, 2, 2)$	$p_2 = -(1/4)\alpha^3 + (7/32)\alpha^2 + (1/32)\alpha$	2
$S_3(1, 1, 1, 2)$	$p_3 = -(1/2)\alpha^3 + (15/32)\alpha + (1/32)$	2
$S_4(1, 1, 2, 3)$	$p_4 = (1/2)\alpha^3 - (23/32)\alpha^2 + (13/64)\alpha + (1/64)$	3
$S_5(1, 2, 2, 2)$	$p_5 = -(1/2)\alpha^3 + (15/32)\alpha + (1/32)$	2
$S_6(1, 2, 3, 3)$	$p_6 = (1/2)\alpha^3 - (23/32)\alpha^2 + (13/64)\alpha + (1/64)$	3
$S_7(1, 2, 1, 2)$	$p_7 = -(1/2)\alpha^3 - (1/32)\alpha^2 + (19/64)\alpha + (15/64)$	2
$S_8(1, 2, 1, 3)$	$p_8 = 2\alpha^3 - (31/16)\alpha^2 - (17/32)\alpha + (15/32)$	3
$S_9(1, 2, 3, 4)$	$p_9 = -(3/2)\alpha^3 + (43/16)\alpha^2 - (11/8)\alpha + (3/16)$	4

NOTE.—In the last column, the number N_a of different alleles is given for each state.

fected sib pairs can be derived for various disease parameters, and not only for dominant or recessive models, without tedious computations. Moreover, data on sib pairs sampled from nonpanmictic populations, as well as data concerning affected individuals with different genealogical relationships, can be dealt with.

Application

Consider sib pairs whose parents are first-cousins in a population with mean inbreeding coefficient $\alpha = 0.01$. Their IBW-state probabilities are reported in table 5 (for the IB-state probabilities, see also appendix C). In table 6, IB probabilities are given for a recessive disease (table 6*A*, $x = s = 0$) and a dominant disease (table 6*B*, $x = 1$, $s = 0$) with no phenocopies for various values of the DS allele frequency, q. Results in italics in these tables were obtained by Thomson and Bodmer (1977) for pairs of sibs whose parents are not related. A comparison shows that inbreeding leads to a departure in the expected probabilities of IB states under both the recessive and dominant models. The departure is higher for small frequencies q of the DS allele. Ignoring inbreeding can bias the estimate of q, leading to an overestimation.

Test of Linkage with IBW States

Tests of linkage based on sib pairs usually look at the identity of alleles between individuals (IB states). When data on inbred sib pairs are available, it can, in addition, be more efficient to take into account the identity of alleles within individuals (IBW states). In particular, for a rare recessive disease, we expect sibs to have both alleles at the DS locus IBD and information concerning this within-identity could be used to detect linkage. In such a situation, one can compare the observed numbers of pairs in each of the nine IBW states with the expected numbers of pairs, on the assumption of no linkage, by a χ^2 statistic. However, when one looks at the nine IBW states, expected numbers of pairs may be small in each category. For instance, considering pairs of sibs from first-cousin marriages, the expected probabilities of each of the nine IBW states under the null hypothesis of no linkage are given in table 3. In table 5, the IBW-state

Table 4

Conditional Probabilities of Genotypes at the Disease-Susceptibility Locus, Given the Identity State for Two Individuals at a Marker Locus Tightly Linked to That Locus

ID State Genotype	S_1 (1, 1, 1, 1)	S_2 (1, 1, 2, 2)	S_3 (1, 1, 1, 2)	S_4 (1, 1, 2, 3)	S_5 (1, 2, 2, 2)	S_6 (1, 2, 3, 3)	S_7 (1, 2, 1, 2)	S_8 (1, 2, 1, 3)	S_9 (1, 2, 3, 4)
DD, DD	q	q^2	q^2	q^3	q^2	q^3	q^2	q^3	q^4
DD, Dd	0	0	$q(1-q)$	$2q^2(1-q)$	0	0	0	$q^2(1-q)$	$2q^3(1-q)$
DD, dd	0	$q(1-q)$	0	$q(1-q)^2$	0	$q^2(1-q)$	0	0	$q^2(1-q)^2$
Dd, DD	0	0	0	0	$q(1-q)$	$2q^2(1-q)$	0	$q^2(1-q)$	$2q^3(1-q)$
Dd, Dd	0	0	0	0	0	0	$2q(1-q)$	$q(1-q)$	$4q^2(1-q)^2$
Dd, dd	0	0	0	0	$q(1-q)$	$2q(1-q)^2$	0	$q(1-q)^2$	$2q(1-q)^3$
dd, DD	0	$q(1-q)$	0	$q^2(1-q)$	0	$q(1-q)^2$	0	0	$q^2(1-q)^2$
dd, Dd	0	0	$q(1-q)$	$2q(1-q)^2$	0	0	0	$q(1-q)^2$	$2q(1-q)^3$
dd, dd	$1-q$	$(1-q)^2$	$(1-q)^2$	$(1-q)^3$	$(1-q)^2$	$(1-q)^3$	$(1-q)^2$	$(1-q)^3$	$(1-q)^4$

Table 5

Probabilities of IBW States for a Pair of Affected Sibs as a Function of the Parameters at the Disease-Susceptibility Locus (q Frequency of the Susceptibility Allele and Penetrances $x = f_2/f_1$; $s = f_3/f_1$)

State	Probability $q_i' = p_i' \times A$
S_1	$p_1[q + (1 - q)s^2]$
S_2	$p_2[q^2 + 2q(1 - q)s + (1 - q)^2s^2]$
S_3	$p_3[q^2 + q(1 - q)(x + s) + (1 - q)^2s^2]$
S_4	$p_4[q^3 + 2q^2(1 - q)x + q(1 - q)s + 2q(1 - q)^2xs + (1 - q)^3s^2]$
S_5	$p_5[q^2 + q(1 - q)(x + s) + (1 - q)^2s^2]$
S_6	$p_6[q^3 + 2q^2(1 - q)x + q(1 - q)s + 2q(1 - q)^2xs + (1 - q)^3s^2]$
S_7	$p_7[q^2 + 2q(1 - q)x^2 + (1 - q)^2s^2]$
S_8	$p_8[q^3 + 2q^2(1 - q)x + q(1 - q)x^2 + 2q(1 - q)^2xs + (1 - q)^3s^2]$
S_9	$p_9[q^4 + 4q^3(1 - q)x + 2q^2(1 - q)^2s + 4q^2(1 - q)^2x^2 + 4q(1 - q)^3xs + (1 - q)^4s^2]$

NOTE.—For sib pairs from first-cousin matings, the probabilities p_i (i = 1,9) are those reported in table 3. The factor A is the probability for the sib pair to be affected: $A = \Sigma_i\, q_i'$.

probabilities are given under the alternative hypothesis of linkage as a function of the genetic model at the DS locus. The states S_3 and S_5, as well as S_4 and S_6, always have the same probabilities when sibs of first-cousins are considered and thus can be pooled. Our purpose is to find an efficient pooling scheme for the nine IBW states. Instead of categorizing on the basis of the number of alleles shared IBD by the two sibs, we propose to categorize the data on the basis of the total number of different alleles N_a (see last row of the table 3). The following categories will be considered:

- Category 1: N_a=1: IBW state S_1
- Category 2: N_a =2: IBW states S_2, S_3, S_5, and S_7
- Category 3: N_a = 3: IBW states S_4, S_6, and S_8
- Category 4: N_a = 4: IBW state S_9

The test of linkage will consist in comparing the observed number of pairs in each of these categories with the expected number under the null hypothesis of no linkage. Under H_0 (no linkage) and for large sample sizes, the statistic follows a χ^2 distribution with 3 df. This test is denoted the "N_a test." This categorization is not the only one possible and different ones could be proposed depending on the DS model. For a rare recessive disease, we expect sib pairs to be in the category 1 (N_a = 1) more frequently than expected at random, whereas for a dominant disease, we expect the third category (N_a = 3) to be more frequent.

To compare the power of our N_a test to the power of the IB test, we determined analytically for a sample of N sib pairs from first-cousin marriages (from a population with mean inbreeding coefficient $\alpha = 0.01$) the expected probabilities in the N_a and IB categories under hypothesis H_1 of linkage and genetic model G_m at the DS locus (see appendix C). Under H_1, the IB and the N_a statistics follow a noncentral χ^2 with, respectively, 2 and 3 df and with noncentrality parameters λ_{IB} and λ_{N_a} that depend on the model G_m. We computed the noncentrality parameters for different models at the DS locus and determined the power of the two tests for these models. The power was studied as a function of the values of the relative penetrances, the number of sib pairs N and the frequency of the DS allele.

In figure 2, the power for 30 sib pairs is given as a function of x, the ratio of the penetrances f_2/f_1 (no phenocopies are assumed: $f_3 = 0$) and assuming a DS allele frequency q of 0.2. For $x < 0.5$, the N_a test appears to be the more powerful of the two tests. For $x = 0.2$, the power of the N_a test is ~0.95, whereas it is ~0.65 for the IB test. This gain in power of ~25% may not be negligible in terms of the number of sib pairs necessary to detect linkage. For $x = 0$, that is to say, for a recessive model with no phenocopies ($f_1 \neq 0$; $f_2 = f_3 = 0$), the gain in power is not perceptible, since the power of the three tests is approximately equal to 1 for the sample sizes considered. In this figure, the power of the IBW test where six categories are considered: S_1, $S_3 + S_5$, $S_4 + S_6$, S_7, S_8, and S_9 are also reported. We can see that the pooling of categories in the N_a test does not lead to a significant loss of power. For $x \leq 0.8$, the N_a test even has a slightly greater power than the IBW test.

We have reported in figure 3 the power of the two tests (N_a and IB) when sporadic cases are allowed for (N = 30 sib pairs). The penetrance f_2 was set to $0.2f_1$ and f_3 to $s \times f_1$, with s varying from 0 to 0.2 (q = 0.2). The test based on N_a is the more powerful of the two tests, whatever s. The 30% gain in power appears to be almost constant with respect to s.

Results are also reported for different sample sizes N, considering a DS allele frequency q of 0.2 and the penetrances f_2 and f_3 equal to $0.2f_1$ and $0.1f_1$, respectively. It can be seen in figure 4 that, on average, 40 sib pairs are required for the power of the N_a test to exceed 0.80 and that this power is not reached with 100 sib

Table 6

Probabilities for an Affected Sib Pair from First-Cousin Marriages (in a population with α = .01) to Share Both (*X*), One (*Y*), or No (*Z*) Alleles Identical by Descent for Various Values of the Disease-Susceptibility Allele Frequency *q* in Recessive and Dominant Cases

q	*X*	*Y*	*Z*
	A. Recessive Cases		
.00	1.000 *1.000*	.000 *.000*	.000 *.000*
.05	.859 *.907*	.137 *.091*	.004 *.002*
.10	.771 *.827*	.219 *.165*	.010 *.008*
.15	.703 *.756*	.278 *.227*	.019 *.017*
.20	.647 *.694*	.325 *.278*	.028 *.028*
.25	.598 *.640*	.363 *.320*	.039 *.040*
.30	.556 *.592*	.394 *.355*	.050 *.053*
.35	.519 *.549*	.420 *.384*	.061 *.067*
.40	.485 *.510*	.442 *.408*	.073 *.082*
.45	.456 *.476*	.460 *.428*	.084 *.096*
.50	.428 *.445*	.476 *.444*	.096 *.111*
.55	.404 *.416*	.488 *.458*	.108 *.126*
.60	.381 *.391*	.499 *.469*	.120 *.140*
.65	.361 *.367*	.508 *.478*	.131 *.155*
.70	.342 *.346*	.515 *.484*	.143 *.170*
.75	.324 *.326*	.521 *.490*	.155 *.184*
.80	.309 *.309*	.525 *.494*	.166 *.198*
.85	.294 *.292*	.529 *.497*	.177 *.211*
.90	.280 *.277*	.532 *.499*	.188 *.224*
.95	.267 *.263*	.534 *.500*	.199 *.237*
1.00	.255 *.250*	.535 *.500*	.210 *.250*
	B. Dominant Cases		
.00	.500 *.500*	.500 *.500*	.000 *.000*
.05	.448 *.460*	.518 *.495*	.034 *.045*
.10	.421 *.428*	.517 *.491*	.062 *.081*
.15	.397 *.401*	.517 *.488*	.086 *.111*
.20	.377 *.377*	.517 *.487*	.106 *.136*
.25	.360 *.358*	.517 *.486*	.123 *.156*
.30	.345 *.341*	.517 *.485*	.138 *.174*
.35	.331 *.326*	.518 *.485*	.151 *.189*
.40	.320 *.313*	.519 *.486*	.161 *.201*
.45	.309 *.302*	.520 *.487*	.171 *.211*
.50	.300 *.293*	.522 *.488*	.178 *.219*
.55	.292 *.284*	.523 *.489*	.185 *.227*
.60	.285 *.277*	.525 *.490*	.190 *.323*
.65	.278 *.271*	.527 *.492*	.195 *.237*
.70	.273 *.265*	.528 *.494*	.199 *.241*
.75	.268 *.261*	.530 *.495*	.202 *.244*
.80	.264 *.257*	.531 *.497*	.205 *.246*
.85	.261 *.254*	.533 *.498*	.206 *.248*
.90	.258 *.252*	.534 *.499*	.208 *.249*
.95	.256 *.250*	.535 *.500*	.209 *.250*
1.00	.255 *.250*	.535 *.500*	.210 *.250*

NOTE.—In italics, results of Thomson and Bodmer (1977) for non-inbred sib pairs.

pairs for the IB test. In this situation, we can expect to obtain a significant result with less than half the number of sib pairs when using the N_a test.

In figure 5, the power is given as a function of the frequency q of the DS allele. The sample size was set to 60 sib pairs (it is the size that leads to a power for the N_a test of ~0.95) and the penetrances were set to f_1, $f_2 = 0.2f_1$, and $f_3 = 0.1f_1$. For this set of penetrances, the power is maximum for the N_a test when q is 0.15 (it is 0.97 for the N_a test and ~0.40 for the IB test). For larger values of q, the power decreases; for instance, if $q = 0.4$, the power is 0.67 for the N_a test and 0.42 for the IB test. The power also decreases for values of q <0.15. This can be explained because the ratio s is constant and thus the proportion of sporadic cases increases with decreasing values of q.

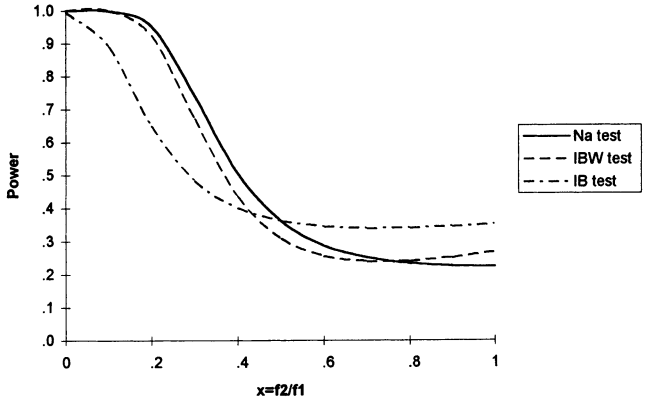

Figure 2 Comparison of the power of the N_a and IB tests for a nominal value of 5% as a function of x, the ratio of the penetrances f_2/f_1. f_3 is set to 0 (no phenocopies) and q to 0.2. Samples of 30 sib pairs from a first-cousin marriage, in a population with mean inbreeding coefficient α = 0.01, are considered. The power is also reported in the case where IBW categories are not pooled (IBW test).

To study the power of tests, we have considered a population with mean inbreeding coefficient α = 0.01. This was, for instance, the average inbreeding coefficient reported in the Egyptian population (Hafez et al. 1983). In figure 6, the power of the two tests has been reported for different levels of remote consanguinity. Power cal-

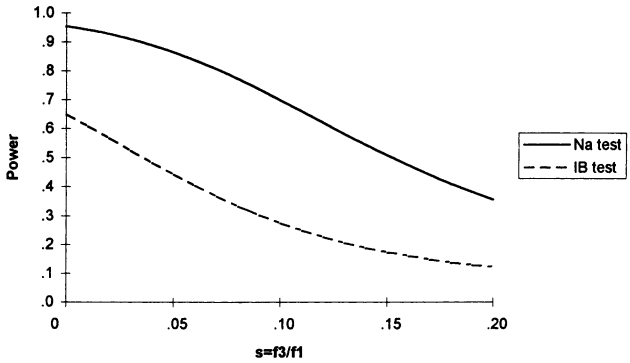

Figure 3 Comparison of the power of the N_a and IB tests for a nominal value of 5% as a function of s, the ratio of the penetrances f_3/f_1. f_2 is set to $0.2f_1$ and q to 0.2.

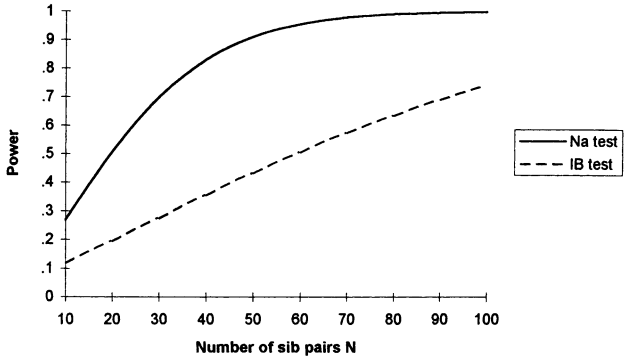

Figure 4 Comparison of the power of the N_a and IB tests for a nominal value of 5% as a function of the size N of the sample of sib pairs. f_2 is set to $0.2f_1$, f_3 to $0.1f_1$, and q to 0.2.

culations have been performed, considering a sample of 30 sib pairs, and the parameters at the DS locus were set to the same values as in figure 4 ($q = 0.2$; $x = 0.2$; and $sp = 0.1$). When the level of remote consanguinity increases, the power of the IB and N_a tests is increased. For the IB test, the increase in power is very small, from 0.26 for $\alpha = 0$ to 0.27 for $\alpha = 0.05$. The increase in power for the N_a test is greater: from 0.67 for $\alpha = 0$ to 0.78 for $\alpha = 0.05$.

Robustness of Tests in Presence of Remote Consanguinity

The presence of a remote consanguinity in the population where sib pairs are sampled will lead to the possibility for the parents of the sibs to be more related than expected by their known first-cousin relationship and to be inbred themselves. This remote consanguinity is usually very difficult to evaluate, and for this reason it is very often ignored. In this section, we studied whether the ignorance of the remote consanguinity may bias the tests. Comparing the distribution of IBW states for sib

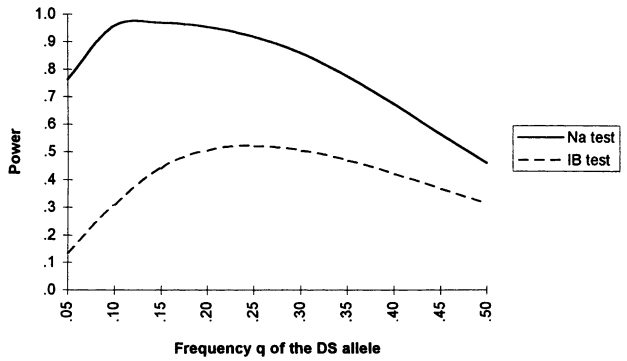

Figure 5 Comparison of the power of the N_a and IB tests for a nominal value of 5% as a function of the frequency q of the DS allele. f_2 is set to $0.2f_1$, f_3 to $0.1f_1$, and the sample size N of sib pairs from a first-cousin marriage is 60.

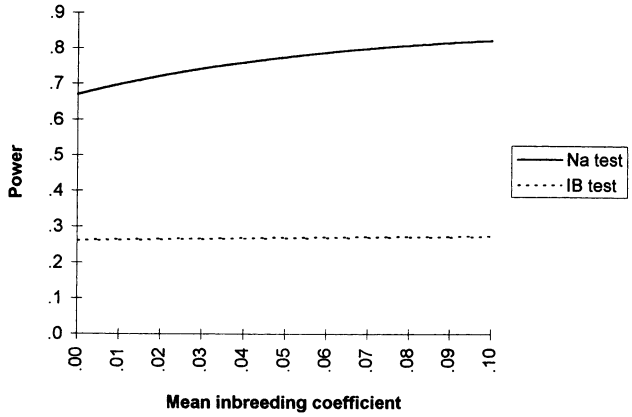

Figure 6 Comparison of the power of the N_a and IB tests as a function of the level of remote consanguinity in the population (mean inbreeding coefficient) for a nominal value of 5%. Power are calculated considering samples of 30 sib pairs and the same model as in figure 4 ($q = 0.2$; $f_2 = 0.2f_1$; $f_3 = 0.1f_1$).

pairs from first-cousin marriages without accounting for the remote consanguinity (H_0) ($\alpha = 0$) and accounting for it ($H_{0\alpha}$), we can see the impact of the remote consanguinity on the IB and the N_a tests. Indeed, since the N_a and the IB statistics are expected to follow a noncentral χ^2 under $H_{0\alpha}$, we were able to compute analytically the type I error of these two tests for samples of N sib pairs from first cousins as a function of α. In figure 7, results are given for a nominal value of 5% for $N = 30$ and $N = 100$ sib pairs. We can see that the N_a test is less robust to the presence of an ignored remote consanguinity than is the IB test. This could be expected, since the N_a test takes into account the within-identity that may be increased by the remote consanguinity. For $\alpha = 0.01$, the type I error is increased from its nominal value of 5% to 5.1% and 5.3% for the IB and the N_a tests, respec-

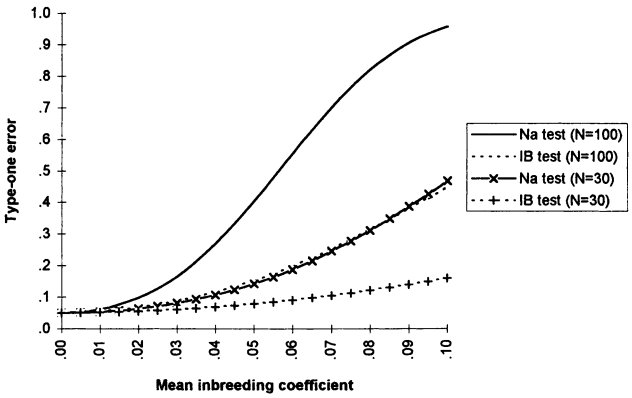

Figure 7 Type I errors of the IB and N_a tests as a function of the level of remote consanguinity in the population (mean inbreeding coefficient) for a nominal value of 5%. The type I error is derived for samples of 100 sib pairs and 30 sib pairs.

tively, if $N = 30$. If the sample size is 100, type I errors are 5.4% and 6.1%, respectively.

If the population is highly consanguineous, the ignorance of the remote consanguinity may increase the type I error of the tests dramatically. For $\alpha = 0.05$ and $N = 30$, the type I error is increased from its nominal value of 5% to 7.8% and 14.3% for the IB and the N_a tests, respectively. For $N = 100$, type I errors become 15.2% and 40.4%, respectively.

Discussion

To test for linkage between a marker locus and a disease, one usually looks for identical marker alleles in related affected individuals. When data on affected children from consanguineous marriages are available, one can also look for regions of their genome that are homozygous by descent (or autozygous) (Lander and Botstein 1987). This method of homozygosity mapping was shown to be very efficient in the study of recessive diseases. We have combined both approaches in this article by using IBW states for affected sib pairs. These IBW states give a complete description of the IBD relationships between alleles of individuals at a locus.

In this paper, we first used the IBW states to derive the expected proportions of sib pairs sharing zero, one, or both marker alleles IBD when sibs are sampled in a consanguineous population. A departure from the proportions of 1/4, 1/2, and 1/4, expected when assuming no inbreeding, was found in a consanguineous population in the absence of linkage. For inbreeding coefficients >0.01, the type I errors of the classical tests of linkage based on sib pairs may largely be greater than their nominal values. Thus, it seems that ignoring inbreeding can lead to a false conclusion of linkage. It may also lead to false estimates of the genetic parameters at the DS locus. In particular, we showed here that the DS allele frequency may be overestimated by using the IBD distribution (Thomson and Bodmer 1977) when inbreeding is not taken into account.

In the second part of the paper, we show how to better use the information provided by data from consanguineous populations. We categorized the data in another way than the usual number of alleles shared IBD between two affected sibs. When concerned with pairs of inbred sibs, we can look at the four alleles of these sibs at a locus and categorize the data on the basis of the number of different alleles N_a in this set of four alleles. We showed that this proposed N_a test is more powerful to detect linkage between a marker and a DS locus than the usual test based on the number of alleles shared IBD by the sibs (IB test) if the ratio x of the penetrances f_2/f_1 is <0.5. The gain in power the N_a test is used has been investigated for different DS allele frequencies q and for different sets of penetrances at the DS locus. The higher power of the N_a test compared to the IB test is especially perceptible for intermediate models where the penetrance of the heterozygote f_2 is ~ 0.2 times that of the homozygote DD, f_1. In the presence of sporadic cases ($f_3 \neq 0$), the N_a test is still the most powerful. In these situations, we can expect to detect linkage with smaller samples of sib pairs from first-cousin marriages. For instance, if $q = 0.2$, $f_2 = 0.2f_1$, and $f_3 = 0.1f_1$, an average sample size of 40 is needed to obtain a power >0.80 for the N_a test, and this sample size is >100 for the IB test and for the IB sib test. The situations where the use of the N_a test may be interesting are situations we can expect when considering multifactorial diseases. Indeed, in such diseases, genetic factors that are not necessarily rare in the population may play a role and interact with each other and with environmental factors.

Although the N_a test is more powerful than the IB test where inbred sib pairs are concerned, it is less robust to the presence of remote consanguinity. If the mean inbreeding coefficient of the population where inbred sibs are sampled is ignored, the type I error of the N_a test may be considerably increased. Accurate estimates of inbreeding coefficients are thus necessary to ensure that type I errors of tests remain close to their nominal value. This is especially true in highly consanguineous populations or subpopulations with mean inbreeding coefficients >0.01. Moreover, in this case, we have shown that accounting for the remote consanguinity can increase the power of the N_a test.

Although the concept of IBW states proposed by Gillois (1964) (who used the terminology "identity states") did not have many practical applications in human genetics, it has been often used in animal genetics, where genealogical relationships are easier to establish thanks to controlled breeding (Chevalet et al. 1984). We show here that, because IBW state give a complete description of the IBD relationships between a set of genes, it may also be a useful tool in genetic epidemiology. It allows us to extend the sib-pair method to data concerning individuals who are not sibs (for instance, a pair of first-cousins) without tedious computations. Indeed, IBW states are a way to deal with complex genealogical relationships and to better take into account genealogical knowledge. Jacquard (1972) used IBW states to derive the genotype structure of an individual given genotypic information concerning one of his relatives. Thompson (1983) used IBW states to infer the ancestral origin of the genes of an individual from a pedigree. In the field of linkage analysis, IBW states have also given rise to some interesting methods. In particular, the affected-pedigree-member method of Weeks and Lange (1988) takes advantage of the IBW states to derive a statistic that permits an investigator to determine whether affected members from a pedigree share more marker alleles IBD than is expected by their genealogical relation-

ships. More recently, Whittemore and Halpern (1994) have also proposed pedigree-linkage tests based on the IBW states. These examples illustrate that the computation of the probabilities of IBW states may be a major step in the analysis of pedigree data. However, neither the approach of Weeks and Lange (1988) nor the approach of Whittemore and Halpern (1994) takes full advantage of IBW states, since they only use a part of the information that concerns IBD between individuals. Contrary to our approach, no attention is given to the IBD within individual, although this may be an important information, as illustrated by our study.

Acknowledgments

We would like to thank Heike Bickeböller and Marie-Claude Babron for their fruitful comments on this manuscript. Special thanks to Bernard Prum for his help with noncentral χ^2's and to the reviewers for their very helpful comments.

Appendix A

IBW-State Probabilities for Two Sibs in a Consanguineous Population

Let us consider that the population from which sib pairs are sampled is consanguineous with mean inbreeding coefficient α (probability to be IBD for the two alleles at a locus of a randomly selected individual in the population). We assume that this population is stable so that the mean inbreeding coefficient stays the same across generations. The mean kinship coefficient in this population (probability to be IBD for two alleles taken at random at a locus from two randomly selected individuals in the population) is then also α. Thus, α is the probability to be IBD for two alleles taken at random in the population (either in the same or in two different individuals).

Let consider the parents of the sib pair. It is possible to compute their probability to be in one of the nine IBW states as a function of α.

S_1 (1 1 1 1) α^3

S_2 (1 1 2 2) $\alpha^2(1-\alpha)$

S_3 (1 1 1 2) $\alpha^2(1-\alpha)$

S_4 (1 1 2 3) $\alpha(1-\alpha)^2$

S_5 (1 2 2 2) $\alpha^2(1-\alpha)$

S_6 (1 2 3 3) $\alpha(1-\alpha)^2$

S_7 (1 2 1 2) $\alpha^2(1-\alpha)^2$

S_8 (1 2 1 3) $2\alpha(1-\alpha)^3$

S_9 (1 2 3 4) $(1-\alpha)^4$

Table A1

IBW-State Probabilities of Parents and Sib Pairs

IBW States of Sibs	IBW State of Parents								
	S1	S2	S3	S4	S5	S6	S7	S8	S9
S1	1	0	1/4	0	1/4	0	1/8	1/16	0
S2	0	0	0	0	0	0	1/8	0	0
S3	0	0	1/4	0	1/4	0	1/4	1/8	0
S4	0	0	0	0	0	0	0	1/16	0
S5	0	0	1/4	0	1/4	0	1/4	1/8	0
S6	0	0	0	0	0	0	0	1/16	0
S7	0	1	1/4	1/2	1/4	1/2	1/4	3/16	1/4
S8	0	0	0	1/2	0	1/2	0	3/8	1/2
S9	0	0	0	0	0	0	0	0	1/4

Once the IBW-state probabilities of parents are known, the IBW-state probabilities for the sib pair can be obtained by using the matrix M_{ps} in table A1.

The IBW-state probabilities for the sib pair are thus:

S_1 (1 1 1 1) $(5/8)\alpha^3 + (1/4)\alpha^2 + (1/8)\alpha$

S_2 (1 1 2 2) $(1/8)\alpha^4 - (1/4)\alpha^3 + (1/8)\alpha^2$

S_3 (1 1 1 2) $-(1/4)\alpha^3 + (1/4)\alpha$

S_4 (1 1 2 3) $-(1/8)\alpha^4 + (3/8)\alpha^3 - (3/8)\alpha^2 + (1/8)\alpha$

S_5 (1 2 2 2) $-(1/4)\alpha^3 + (1/4)\alpha$

S_6 (1 2 3 3) $-(1/8)\alpha^4 + (3/8)\alpha^3 - (3/8)\alpha^2 + (1/8)\alpha$

S_7 (1 2 1 2) $(1/8)\alpha^4 - (7/8)\alpha^3 + (1/8)\alpha^2 + (3/8)\alpha + (1/4)$

S_8 (1 2 1 3) $-(1/4)\alpha^4 + (5/4)\alpha^3 - (5/4)\alpha^2 - (1/4)\alpha + (1/2)$

S_9 (1 2 3 4) $(1/4)\alpha^4 - \alpha^3 + (3/2)\alpha^2 - \alpha + (1/4)$

Appendix B

Accounting for Remote Consanguinity in the Computation of IBW-State Probabilities for Two Sibs from First-Cousin Matings

IBW-state probabilities can be computed on the basis of genealogical information by using the algorithm of Karigl (1981). In this algorithm, it is assumed that individuals outside the pedigree are neither related nor inbred. Thus, it is assumed that the mean inbreeding and the mean kinship coefficient are negligible in the population. Since this may not be the case in populations where inbred sib pairs may be collected, we try to take into account this remote consanguinity. As in appendix A, we suppose that the population is characterized by a mean inbreeding coefficient equal to the mean kinship coefficient and equal to α. We introduce this mean inbreeding coefficient in the computation of IBW states for first cousins, using the algorithm of Karigl (1981).

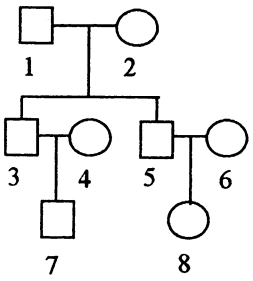

Figure B1 Pedigree, showing individuals (7 and 8) as first cousins, to compute IBW states.

Let consider the following pedigree in figure B1, where individuals 7 and 8 are first cousins. To deduce the nine IBW states, Karigl (1981) introduced generalized kinship coefficients. Six kinship coefficients, ϕ, need to be computed to obtained the IBW-state probabilities of individuals 7 and 8. These kinship coefficients can be obtained by using the recursion rules given by Karigl (1981).

$$\begin{aligned}\phi_{78} &= \phi_{87}\\ &= (1/4)(\phi_{35} + \phi_{36} + \phi_{45} + \phi_{46})\\ &= (1/4)(\phi_{35} + 3\alpha),\\ &\quad \text{since} \quad \phi_{36} = \phi_{45} = \phi_{46} = \alpha\\ &= (1/16)(\phi_{11} + \phi_{12} + \phi_{21} + \phi_{22}) + (3/4)\alpha\\ &= (1/16)[(1/2)(1 + \alpha) + 2\alpha + (1/2)(1 + \alpha)]\\ &= (1/16) + (15/16)\alpha\end{aligned}$$

$$\phi_{77} = \phi_{88} = (1/2)(1 + \phi_{34}) = (1/2)(1 + \alpha)$$

$$\begin{aligned}\phi_{778} &= \phi_{887}\\ &= (1/2)(\phi_{78} + \phi_{348})\\ &= (1/2)\phi_{78} + (1/4)(\phi_{345} + \phi_{346})\\ &= (1/2)\phi_{78} + (1/4)[(1/2)(\phi_{145} + \phi_{245}) + \alpha^2]\\ &= (1/2)\phi_{78} + (1/4)[(1/4)(\phi_{115} + \phi_{125}\\ &\quad + \phi_{214} + \phi_{224}) + \alpha^2]\\ &= (1/2)\phi_{78} + (1/4)[(1/4)(4\alpha^2) + \alpha^2]\\ &= (1/32) + (15/32)\alpha + (1/2)\alpha^2\end{aligned}$$

$$\begin{aligned}\phi_{7788} &= (1/2)(\phi_{788} + \phi_{3488})\\ &= (1/2)\phi_{788} + (1/4)(\phi_{348} + \phi_{3456})\\ &= (1/2)\phi_{788} + (1/4)\phi_{348}\\ &\quad + (1/4)(1/2)(\phi_{1456} + \phi_{2456})\\ &= (1/2)\phi_{788} + (1/4)\phi_{348}\\ &\quad + (1/8)(1/2)(\phi_{1416} + \phi_{1426} + \phi_{2416} + \phi_{2426})\\ &= (1/64) + (15/64)\alpha + (1/4)\alpha^2\\ &\quad + (1/4)\alpha^2 + (1/16)(4\alpha^3)\\ &= (1/64) + (15/64)\alpha + (1/2)\alpha^2 + (1/4)\alpha^3\end{aligned}$$

$$\begin{aligned}\phi_{77,88} &= 2(1/2)(\phi_{88} + \phi_{34,88})\\ &= (1/2)\phi_{88} + (1/2)(1/2)(\phi_{34} + \phi_{34,56})\\ &= (1/2)\phi_{88} + (1/4)\phi_{34} + (1/4)\alpha^2\\ &= (1/4)(1 + \alpha) + (1/4)\alpha + (1/4)\alpha^2\\ &= (1/4) + (1/2)\alpha + (1/4)\alpha^2\end{aligned}$$

$$\begin{aligned}\phi_{78,78} &= (1/4)(2\phi_{78} + 2\phi_{38,48})\\ &= (1/2)\phi_{78} + (1/2)(1/4)(2\phi_{348} + \phi_{53,64} + \phi_{63,54})\\ &= (1/2)\phi_{78} + (1/4)\phi_{348} + (1/8)\phi_{53,64} + (1/8)\alpha^2\\ &= (1/2)\phi_{78} + (1/4)\phi_{348}\\ &\quad + (1/8)(1/4)(\phi_{11,64} + \phi_{12,64} + \phi_{21,64} + \phi_{22,64})\\ &\quad + (1/8)\alpha^2\\ &= (1/64) + (15/64)\alpha + (1/4)\alpha^2 + (1/4)\alpha^2\\ &\quad + (1/32)[2(1/2)(\alpha + \alpha^2) + 2\alpha^2] + (1/8)\alpha^2\\ &= (1/64) + (17.64)\alpha + (23/32)\alpha^2\,.\end{aligned}$$

The IBW-state probabilities for individuals 7 and 8 can be deducted from these extended kinship coefficients as shown by Karigl (1981).

For first cousins in a population with mean inbreeding coefficient α, IBW states have probabilities

$$\mathrm{P(S1)} = 4\phi_{7788} - 2\phi_{778} - 2\phi_{788} + \phi_{78} = \alpha^3$$

$$\begin{aligned}\mathrm{P(S2)} &= -4\phi_{7788} + 4\phi_{77,88} + 2\phi_{778} - 2\phi_{77} + 2\phi_{788}\\ &\quad - 2\phi_{88} - \phi_{78}\\ &= -\alpha^3 + \alpha^2\end{aligned}$$

$$\begin{aligned}\mathrm{P(S3)} &= -8\phi_{7788} + 8\phi_{778} + 4\phi_{788} - 4\phi_{77}\\ &= -2\alpha^3 + 2\alpha^2\end{aligned}$$

$$\begin{aligned}\mathrm{P(S4)} &= 8\phi_{7788} - 4\phi_{77,88} - 8\phi_{778} + 4\phi_{78}\\ &\quad - 4\phi_{788} + 2\phi_{88} + 4\phi_{78} - 2\\ &= 2\alpha 3 - 3\alpha^2 + \alpha\end{aligned}$$

$$\begin{aligned}\mathrm{P(S5)} &= -8\phi_{7788} + 4\phi_{778} + 8\phi_{788} - 4\phi_{78}\\ &= -2\alpha^3 + 2\alpha^2\end{aligned}$$

$$\mathrm{P(S6)} = 8\phi_{7788} - 4\phi_{77,88} - 4\phi_{778} + 2\phi_{77}$$

$$
\begin{aligned}
&- 8\phi_{788} + 4\phi_{88} + 4\phi_{78} - 2 \\
&= 2\alpha^3 - 3\alpha^2 + \alpha
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{P(S7)} &= -8\phi_{7788} + 8\phi_{78,78} \\
&= -2\alpha^3 + (7/4)\alpha^2 + (1/4)\alpha
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{P(S8)} &= 32\phi_{7788} - 16\phi_{778} - 16\phi_{788} \\
&\quad - 16\phi_{78,78} + 16\phi_{78} \\
&= 8\alpha^3 - (23/2)\alpha^2 + (13/4)\alpha + (1/4)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{P(S9)} &= -24\phi_{7788} + 4\phi_{77,88} + 16\phi_{778} - 4\phi_{77} \\
&\quad + 16\phi_{788} - 4\phi_{88} + 8\phi_{78,78} - 16\phi_{78} + 4 \\
&= -6\alpha^3 + (43/4)\alpha^2 - (11/2)\alpha + (3/4)\,.
\end{aligned}
$$

The use of the matrix $\mathbf{M}_{ps}$, defined in appendix A, allows derivation of the IBW-state probabilities for sib pairs from first-cousin marriages in this population:

S_1 (1 1 1 1) $(1/4)\alpha^3 + (1/2)\alpha^2 + (15/64)\alpha + (1/64)$

S_2 (1 1 2 2) $-(1/4)\alpha^3 + (7/32)\alpha^2 + (1/32)\alpha$

S_3 (1 1 1 2) $-(1/2)\alpha^3 + (15/32)\alpha + (1/32)$

S_4 (1 1 2 3) $(1/2)\alpha^3 - (23/32)\alpha^2 + (13/64)\alpha + (1/64)$

S_5 (1 2 2 2) $-(1/2)\alpha^3 + (15/32)\alpha + (1/32)$

S_6 (1 2 3 3) $(1/2)\alpha^3 - (23/32)\alpha^2 + (13/64)\alpha + (1/64)$

S_7 (1 2 1 2) $-(1/2)\alpha^3 - (1/32)\alpha^2 + (19/64)\alpha + (15/64)$

S_8 (1 2 1 3) $2\alpha^3 - (31/16)\alpha^2 - (17/32)\alpha + (15/32)$

S_9 (1 2 3 4) $-(3/2)\alpha^3 + (43/16)\alpha^2 - (11/8)\alpha + (3/16)$

Appendix C

IB and N_a States Distribution for Affected Sib Pairs under the Hypothesis H_1 of Linkage between the Marker and the DS Locus (θ is assumed to be 0)

In table 5, the IBW-state probabilities have been given as a function of q (DS allele frequency) and of the ratios of penetrances ($x = f_2/f_1$ and $s = f_3/f_1$) for sib pairs. By using this table, it is possible to determine the expected distribution of IB and N_a states, as shown in table C1, where

$$
\begin{aligned}
A &= q(p_1 + q(p_2 + p_3 + p_5 + p_7) \\
&\quad + q^2(p_4 + p_6 + p_8) + q^3 p_9) + q(1 - q) \\
&\quad \times \begin{bmatrix} x(p_3 + p_5 + 2q(p_4 + p_6 + p_8) + 4q^2 p_9) + s(p_4 + p_6 + 2p_7 + 4q(1 - q)p_9) \\ +xs(p_3 + p_5 + 2(1 - q)(p_4 + p_6 + p_8) + 4(1 - q)^2 p_9) \\ +x^2(2p_7 + p_8 + 4q(1 - q)p_9) \end{bmatrix} \\
&\quad + (1 - q)s^2(p_1 + (1 - q)(p_2 + p_3 + p_5 + p_7) \\
&\quad + (1 - q)^2(p_4 + p_6 + p_8) + (1 - q)^3 p_9)\,.
\end{aligned}
$$

By using the probabilities of IBW states (p_i, $i = 1$–9) given in table 3, the IB-state distribution may be obtained for sib pairs from first cousins sampled in a population with mean inbreeding coefficient α. For the N_a states, see table C2.

Under the hypothesis H_1, the N_a and IB statistics follow noncentral χ^2's with, respectively, 2 and 3 df. The noncentral parameters λ_{IB} and λ_{N_a} can be computed from the above distribution:

$$
\lambda_{IB} = N \sum_{i=1}^{3} \frac{(p_i - p_i')}{p_i} \quad \text{and} \quad \lambda_{N_a} = N \sum_{i=1}^{4} \frac{(p_i - p_i')}{p_i}\,,
$$

where p_i and p_i' are the probabilities of state i under H_0 and H_1, respectively.

Table C1

Expected Distribution of IB States

IB State	Under H_0	Under H_1
IB = 0	$p_2 + p_4 + p_6 + p_9$ $\frac{7}{32} - \frac{30}{32}\alpha + \frac{47}{32}\alpha^2 - \frac{24}{32}\alpha^3$	$\frac{1}{A}\begin{bmatrix} q^2(p_2 + q(p_4 + p_6) + q^2p_9 \\ + q(1-q)\begin{pmatrix} 2qx(p_4 + p_6 + 2qp_9) \\ + s(2p_2 + p_4 + p_6 + 2q(1-q)p_9) \\ + 2(1-q)xs(p_4 + p_6 + 2(1-q)p_9) \\ + 4q(1-q)x^2p_9 \end{pmatrix} \\ + (1-q)^2s^2(p_2 + (1-q)(p_4 + p_6) + (1-q)^2p_9) \end{bmatrix}$
IB = 1	$p_3 + p_5 + p_8$ $\frac{17}{32} + \frac{13}{32}\alpha - \frac{62}{32}\alpha^2 + \alpha^3$	$\frac{1}{A}\begin{bmatrix} q^2(p_3 + p_5 + qp_8) \\ + q(1-q)\begin{pmatrix} x(p_3 + p_5 + 2qp_8) \\ + xs(p_3 + p^5 + 2(1-q)p_8) \\ + x^2p_8 \end{pmatrix} \\ + (1-q)^2s^2(p_3 + p_5 + (1-q)p_8) \end{bmatrix}$
IB = 2	$p_1 + p_7$ $\frac{8}{32} + \frac{17}{32}\alpha + \frac{15}{32}\alpha^2 - \frac{8}{32}\alpha^3$	$\frac{1}{A}\begin{bmatrix} q(p_1 + qp_7) \\ + 2q(1-q)x^2p_7 + (1-q)s^2(p_1 + (1-q)p_7) \end{bmatrix}$

NOTE.—All cases are sibs from first cousins.

Table C2

Expected Distribution of *Na* States

N_a State	Under H_0	Under H_1
$N_a = 1$	p_1 $\frac{1}{64} + \frac{15}{64}\alpha + \frac{1}{2}\alpha^2 + \frac{1}{4}\alpha^3$	$\frac{1}{A}[q + (1-q)s^2]p_1$
$N_a = 2$	$p_2 + p_3 + p_5 + p_7$ $\frac{19}{64} + \frac{81}{64}\alpha + \frac{3}{16}\alpha^2 - \frac{7}{4}\alpha^3$	$\frac{1}{A}\begin{bmatrix} q^2(p_2 + p_3 + p_5 + p_7) \\ + q(1-q)(x(1+s)(p_3 + p_5) + 2p_2s + 2p_7x^2) \\ + (1-q)^2s^2(p_2 + p_3 + p_5 + p_7) \end{bmatrix}$
$N_a = 3$	$p_4 + p_6 + p_8$ $\frac{32}{64} - \frac{1}{8}\alpha - \frac{27}{8}\alpha^2 + 3\alpha^3$	$\frac{1}{A}\begin{bmatrix} q^3(p_4 + p_6 + p_8) \\ + q(1-q)\begin{pmatrix} 2xq(p_4 + p_6 + p_8) + s(p_4 + p_6) \\ + 2xs(1-q)(p_4 + p_6 + p_8) + x^2p_8 \end{pmatrix} \\ + (1-q)^2s^2(p_4 + p_6 + p_8) \end{bmatrix}$
$N_a = 4$	p_9 $\frac{12}{64} - \frac{11}{8}\alpha + \frac{43}{16}\alpha^2 - \frac{3}{2}\alpha^3$	$\frac{1}{A}\begin{bmatrix} q^4 + 2q(1-q)\begin{pmatrix} 2q^2x + q(1-q)s \\ + 2q(1-q)x^2 + 2(1-q)^2xs \end{pmatrix} \\ + (1-q)^4s^2 \end{bmatrix}p_9$

NOTE.—All cases are sibs from first cousins.

References

Blackwelder WC, Elston RC (1985) A comparison of sib-pair linkage tests for disease susceptibility loci. Genet Epidemiol 2:85–97

Chevalet C, Gillois M, Vu Tien Khang J (1984) Conditional probabilities of identity of genes at a locus linked to a marker. Genet Sel Evol 16:431–444

Cudworth AG, Woodrow JC (1975) Evidence for HLA linked genes in "juvenil" diabetes mellitus. Br Med J 3:133–135

De Vries RRP, Fat RFM, Lai A, Nijenhuis LE, Van Rood JJ (1976) HLA-linked genetic control of host response to mycobacterium Leprae. Lancet 2:1328–1330

Day NE, Simons MJ (1976) Disease susceptibility genes: their identification by multiple case family studies. Tissue Antigens 8:109–119

Gillois M (1964) La relation d'identité en génétique. Thèse Faculté des Sciences de Paris, Paris

Haseman JK, Elston RC (1972) The investigation of linkage between a quantitative trait and a marker locus. Behav Genet 2:3–19

Hafez M, El-Tahan H, Awadalla M, El-Khayat H, Abdel-Gafar A, Ghoneim M (1983) Consanguineous matings in the Egyptian population. J Med Genet 20:58–60

Jacquard A (1970) Structures génétiques des populations. Masson et Cie, Paris

——— (1972) Genetic information given by a relative. Biometrics 28:1101–1114

Karigl G (1981) A recursive algorithm for the calculation of identity coefficients. Ann Hum Genet 45:299–305

——— (1982) Multiple genetic relationships: joint and conditional genotype probabilities. Ann Hum Genet 46:83-92

Lander ES, Botstein D (1987) Homozygosity mapping: a way to map human recessive trait with the DNA of inbred children. Science 236:1567–1570

Lange K (1986) The affected-sib-pair method using identity by state relations. Am J Hum Genet 39:148–150

Penrose LS (1935) The detection of autosomal linkage in data which consist of pairs of brothers and sisters. Ann Eugen 6:133–138

Suarez BK (1978) The affected sib pair IBD distribution for HLA-linked disease susceptibility genes. Tissue Antigens 12: 87–93

Thomson G, Bodmer WF (1977) The genetics of HLA and disease associations. In: Christiansen FB, Fenchel T (eds) Measuring selection in natural populations. Springer, Berlin, pp 545–564

Thompson EA (1974) Gene identities and multiple relationships. Biometrics 30:667–680

——— (1983) A recursive algorithm for inferring gene origins. Ann Hum Genet 47:143–152

Weeks DE, Lange K (1988) The affected-pedigree-member method of linkage analysis. Am J Hum Genet 42:315–326

Whittemore AS, Halpern J (1994) A class of tests for linkage using affected pedigree members. Biometrics 50: 118–127