Genetic Epidemiology 26: 116–124 (2004)

Power of Direct vs. Indirect Haplotyping in Association Studies

Stuart Thomas,[1,2] David Porteous,[2] and Peter M. Visscher[1]*

[1]*Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh, UK*
[2]*Medical Genetics Section, Molecular Medicine Centre, Western General Hospital, Edinburgh, UK*

Haplotype analysis is essential to studies of the genetic factors underlying human disease, but requires a large sample size of phase-known data. Recently, directly haplotyping individuals was suggested as a means of maximizing the phase-known data from a sample. Haplotyping, however, is much more labor-intensive than indirectly inferring haplotypes from genotypes (genotyping). This study uses simulations to compare the power of each methodology to detect associations between a haplotype and a trait or disease locus under conditions of varying linkage disequilibrium. The relative power of haplotyping over genotyping in association studies increases with decreasing sample size, decreasing linkage disequilibrium, decreasing numbers of marker loci, and decreasing numbers of different haplotypes. In addition, the frequency of the haplotype of interest and the magnitude of its association with the disease affect the power. From a cost-benefit standpoint, genotyping would be favored with large multiplicative risks (relative risk of haplotype >2.5). If case numbers are limiting rather than cost, haplotyping would maximize the information obtained. At small haplotype frequencies (e.g., <0.05), haplotyping is relatively more efficient, but there is little absolute power to detect associations under either methodology. Given the much larger laboratory resources required for direct haplotyping, genotyping would probably be favored under most conditions, but this must be balanced against the unit costs associated with recruitment and phenotyping. In the context of multipurpose, prospective cohort studies (e.g., the UK Biobank study), there may be a general value in establishing a series of directly haplotyped individuals to serve as controls for a number of alternative studies. *Genet Epidemiol* 26:116–124, 2004. © 2004 Wiley-Liss, Inc.

Key words: expectation-maximization; haplotype; genotype; linkage disequilibrium; resource implications.

Grant sponsor: UK Medical Research Council; Grant sponsor: Cunningham Trust; Grant sponsor: UK Biotechnology and Biological Sciences Research Council.
*Correspondence to: Peter Visscher, Institute of Cell, Animal and Population Biology. University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK. E-mail: peter.visscher@ed.ac.uk
Received 22 April 2003; Accepted 5 August 2003
Published online in Wiley InterScience (www.interscience.wiley.com)
DOI: 10.1002/gepi.10300

INTRODUCTION

Genetic association studies of some human diseases indicate that an individual's susceptibility is influenced by relatively few genomic regions, or genes, that confer high individual risk (e.g., *APOE* in late-onset Alzheimer's disease [Corder et al., 1993; Saunders et al., 1993]. In other diseases, susceptibility is conjectured to be influenced by a larger number of regions, each conferring lower individual risk. New approaches to data acquisition and handling are required in order to maximize the efficiency of studies into these subtler associations.

Haplotype analyses are central to many association studies. They facilitate the fine-scale mapping of susceptibility genes through linkage disequilibrium analysis of the surrounding markers [Guo, 1997; Rannala and Slatkin, 1998], and allow inferences to be made about the associations between particular haplotypes and the disease. Central to such studies is the need for phase-known information. Thus, studies are limited by the difficulty and cost of obtaining large enough samples of phase-known data to detect subtle genetic contributions. Several approaches have been used to infer phase information: 1) pedigree analysis; 2) molecular haplotyping, which is limited to short sequences and uses chromosome isolation and long-range PCR [Michalatos-Beloin et al., 1996]; and 3) computer-based algorithms, e.g., parsimony [Clark, 1990], expectation-maximization (EM) [e.g., Excoffier and Slatkin, 1995], and Bayesian approaches [Stephens et al., 2001].

More recently, Douglas et al. [2001] proposed the use of conversion techniques [Yan et al., 2000]

as a means to separate chromosome pairs, allowing haplotypes to be determined directly. An important question concerning resource allocation arises when considering the use of conversion or "direct" haplotyping vs. the more indirect methods outlined above. For the convenience of this study, the word *haplotyping* was used to refer to the direct haplotyping approach, and *genotyping* to the indirect estimation of haplotypes from genotypes using EM. While haplotyping inarguably provides more exact information per individual, it is more expensive and labor-intensive, thereby in practice constraining the sample size. Douglas et al. [2001] briefly examined the relative efficiencies of haplotyping over genotyping. They examined the ratio of expected haplotype frequency variances determined by each method, under the simplifying, but unrealistic, condition of linkage equilibrium between markers. They discovered that for larger numbers of markers (when there are many ambiguous genotype classes), direct haplotyping is much more efficient. Schaid [2002] used simulation to extend the analysis of Douglas et al. [2001] to examine haplotyping vs. genotyping under conditions of linkage disequilibrium (LD). He concluded that any advantage of haplotyping decreases as LD increases, although in many situations the conclusions of Douglas et al. [2001] remain valid.

In this study, we use specifically written simulation software to compare the actual power of the two methodologies to detect significant differences between a diseased and a control population with respect to 1) a particular haplotype of interest, and 2) the haplotype frequency distribution of all haplotypes.

METHODS

SIMULATED DATA SETS

In order to select the most appropriate method for simulation of haplotype distributions, the comparative efficiencies of haplotyping vs. genotyping were examined using the method of Schaid [2002]. In this study, the method was extended to cover more of the parameter space, with all possible haplotype combinations being generated systematically (i.e., using all pairwise, tripletwise, etc., disequilibrium coefficients rather than just selected pairwise associations). In addition, other methods of simulation were examined, including a random approach where haplotype frequencies were generated from a number of prespecified (arbitrary) distributions. Finally, published data from the Chromosome 22 Group at the Sanger Institute, obtained from the World Wide Web at http://www.sanger.ac.uk/HGP/Chr22 [Collins et al., 2003], were used to generate real-data haplotype distributions. It was found that simulating distributions using a random uniform distribution obtained relative efficiencies for haplotyping over genotyping that were representative of those real-data simulations when haplotyping was at its most favored (Fig. 1). The random uniform distribution was therefore used in further studies of relative power of haplotyping over genotyping, since if the relative power is

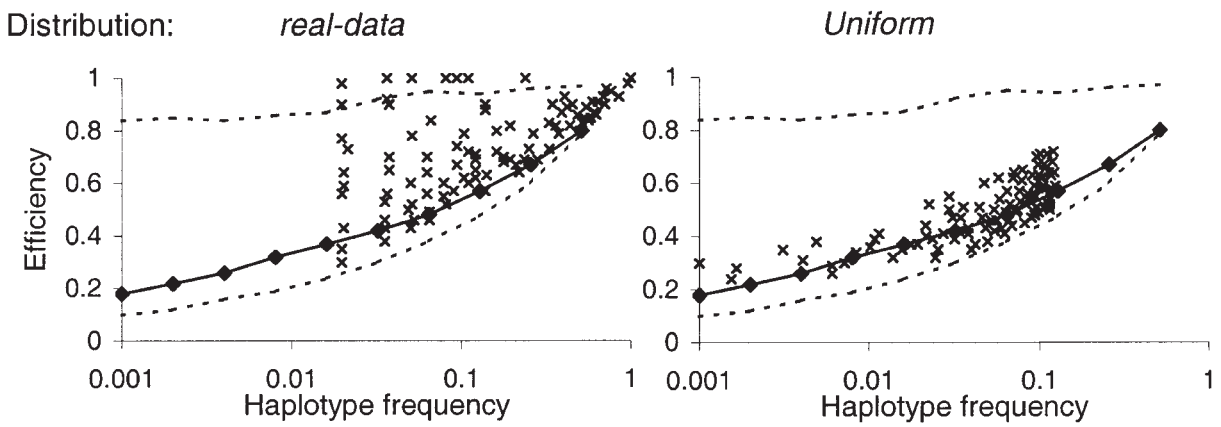

Fig. 1. Relative efficiency of genotyping compared to direct haplotyping for different haplotype frequencies derived from four marker loci. Solid line, average efficiency ratio over haplotype frequency distributions simulated using an extended form of method of Schaid [2002]. Dashed lines, maximum and minimum efficiency values encountered during simulations. Crosses, 100 example points from real data and random uniform methods for simulating haplotype frequencies.

found to be low using this distribution, it will also be low in real-data situations, allowing conclusions to be generalized.

A random uniform distribution was used to generate n_h haplotype frequencies, termed h_i for i=1 to n_h, where $\sum_{i=1}^{n_h} h_i = 1$. The frequency of a disease-associated haplotype (d, with frequency h_d, where d was assigned as 1) was specified in advance, allowing maximum control of the frequency of a haplotype of interest. The remaining n_h-1 haplotype frequencies were generated directly by drawing values from a Uniform(0,1) distribution and scaling them to sum to $1-h_d$.

Individual genotypes were generated by drawing two independent haplotypes from the simulated haplotype frequency distribution. Individuals had a certain baseline risk for a disease, and the presence of d conferred an additive increase, R, to that baseline risk (one copy of d increased risk by R, two copies by $2R$). For the purposes of this study, risk is defined as the additive increase in the probability of an individual having the disease conferred by the presence of d. Based on the simulated disease state, the sample could be split into a ''case'' set containing diseased individuals and a ''control'' set.

HAPLOTYPE FREQUENCY ESTIMATION

Under direct haplotyping, phase is known exactly, and haplotype frequencies are simply governed by a multinomial distribution [e.g., Hill, 1974]. The likelihood of the observed haplotype data was calculated as

$$L_{hap} = (2n)! \prod_{i=1}^{n_h} \frac{(h_{hap,i})^{p_i}}{p_i!}$$

[e.g., Hill, 1974] where p_i is the observed number of haplotypes of type i, and h_{hap} signifies a haplotype frequency estimate made using direct haplotyping. Under genotyping, maximum likelihood haplotype frequency estimates (h_{gen}) are calculated using iterative procedures. In this study, the expectation-maximization (EM) method [Hill, 1974; Dempster et al., 1977; Excoffier and Slatkin, 1995] was used to determine h_{gen}. To ensure that the global maximum for h_{gen} was reached, the EM algorithm specifications determined by Fallin and Schork [2000] were adopted. Their results showed that restarting the EM algorithm 15 times with different initial values for h_{gen}, running each for 150 iterative steps and with convergence criteria set to 10^{-5}, sufficed for all simulations. Assuming a random association of haplotypes, the likelihood of the observed genotype data follows a multinomial distribution and is calculated as:

$$L_{gen} = s_g! \prod_{j}^{n_g} \frac{\pi_j^{g_j}}{g_j!}$$

where g_j is the observed number of type j genotypes; π_j is the expected frequency of the jth genotype calculated from h_{gen} (the frequency of ambiguous genotypes being calculated as the sum of the expected frequency of each of its alternative haplotype combinations); n_g is the number of different genotypes in the sample; and s_g is the total number of genotyped individuals.

Examination was also made of the situation where one part of the sample was genotyped and the other part haplotyped. This required a trivial addition to the EM algorithm, allowing the estimation of haplotype frequencies in the joint sample. This addition involves adding the haplotype counts determined through haplotyping to the maximization step [equation 8, Excoffier and Slatkin, 1995], *before* dividing by twice the number of individuals in the population [i.e., move the division by n from equation 7 to equation 8 in Excoffier and Slatkin, 1995].

POWER STUDIES

Two situations were considered. The first situation was where there is some prior information about d (e.g., from a different study population), thus allowing a ''haplotype-specific'' score test to be used [Schaid et al., 2002]. In brief, the score test uses linear regression to determine the association between d and the disease of interest, using the complete (base and disease) sample. Schaid et al. [2002] determined that for a binary trait the haplotype-specific score approximately followed a χ^2 distribution with 1 degree of freedom. The second and more realistic ''naïve'' test was used when looking for evidence of significantly different haplotype frequency distributions between the two groups, before examining the individual haplotypes. This test utilized the likelihood ratio statistic $Z = 2(\ln(L_{Case}) + \ln(L_{Control}) - \ln(L_{Case+Control}))$, where $\ln(L_x)$ is the natural log of the likelihood for group x [for a review of these methods, see Sham, 1998]. For example, $\ln(L_{Case}) = \ln(L_{hap})$ when haplotype frequencies for the diseased sample are estimated using direct haplotyping. Under the null hypothesis of equal haplotype frequencies in the case and control populations, Z follows an approximate χ^2 distribution with

degrees of freedom equal to n_h-1, where n_h is the number of haplotypes in the combined sample.

It is of note that the score tests of Schaid et al. [2002] may also be used to formulate a naive test, i.e., a general test for haplotype frequency differences between cases and controls. However, this requires that the variance of haplotype scores be approximated from the inverse of the second derivative matrix. Variance estimates made in this manner are inaccurate when there is little haplotype information (due to rarity). When larger numbers of marker loci are simulated (e.g., when each haplotype is comprised of five or more loci), the random uniform method for haplotype frequency generation yields large numbers of rare haplotypes. Under these conditions, the score test method was not found to work well, and so the above likelihood ratio test was adopted. The same problem is encountered to a lesser degree when using a score test to test a specific rare haplotype. Other specific tests (e.g., a *t*-test) to compare h_d in the disease and control groups require similar variance estimates, and so the problem is unavoidable.

Power was determined over a range of h_d and a range of R, with each set of conditions being replicated 1,000 times. In addition, the case and control group sizes were varied in size, up to a maximum of 1,000 individuals in each. Power was determined in four situations: 1) when both case and control samples were haplotyped directly (denoted h|h); 2) when cases were haplotyped and controls genotyped (denoted h|g); 3) when cases were genotyped and controls haplotyped (g|h); and 4) when both were genotyped (g|g). Power was compared for the different conditions with respect to m, h_d, n_h, the relative risk, and the sample sizes.

For each set of simulation conditions, power was calculated at $\alpha=0.05$. With the naive test, the differences between all haplotypes in the groups are examined, and the significance levels represent differences in the whole set rather than just between h_d. Thus random fluctuations between case and control frequencies in any of the other haplotypes will lead to false positives with respect to detectable differences in h_d. In order to minimize such false positives, the case for a more stringent significance level could be argued. However, since this is a comparative study, examining relative power, the choice of significance level is more arbitrary. Lower significance values might be required in order to detect smaller effects and associations when sample size is limited.

RESULTS

The specifications for the EM algorithm determined by Fallin and Schork [2000] were found to work well under all the conditions simulated. For convenience, we will address in turn each aspect of the study that affects power.

TYPE 1 ERRORS

For comparative purposes, it is useful to know the expected number of type 1 errors expected under the null hypothesis, when the increased risk (R) conferred by the specific haplotype (d) is equal to zero (i.e., there is no association between haplotype and disease). When both control and disease group were large (n=1,000), type 1 errors for the specific test lay in the range 0.02 (95% CI=0.011–0.029) to 0.09 (95% CI=0.072–0.108), regardless of haplotype frequency, marker number (m), or method used to determine haplotype frequency (results not shown). The naive test was more dependent on the number of marker loci and method of frequency estimation, showing the largest frequency of type 1 errors (0.13) when there were six markers and frequencies were estimated using EM. In other situations, the frequency of type 1 errors was lower than 0.1. With lower sample sizes, error rates became more conservative (<0.05).

HAPLOTYPE-SPECIFIC VS. NAIVE TEST

In general, the power to detect association between d and associated disease is dependent on the risk, the frequency of the disease haplotype, the number of marker loci, and the sample size (n) (Figs. 2–5). An exception to this is with direct haplotype analysis, when an association between d and the disease is tested (i.e., h|h). In this case, power is independent of m (see h|h−hap in Figs. 3, 4) and the simulated distribution of haplotype frequencies (h|h in Fig. 2).

In most cases, the power of h|h to detect associations is greater than g|g, and the power to detect associations when a particular haplotype is specified is greater than the naive case. With higher numbers of marker loci, however, and at low values for h_d (<0.01), the power of testing the specific case when both samples are genotyped can fall below the power of the naive test (Figs. 3,

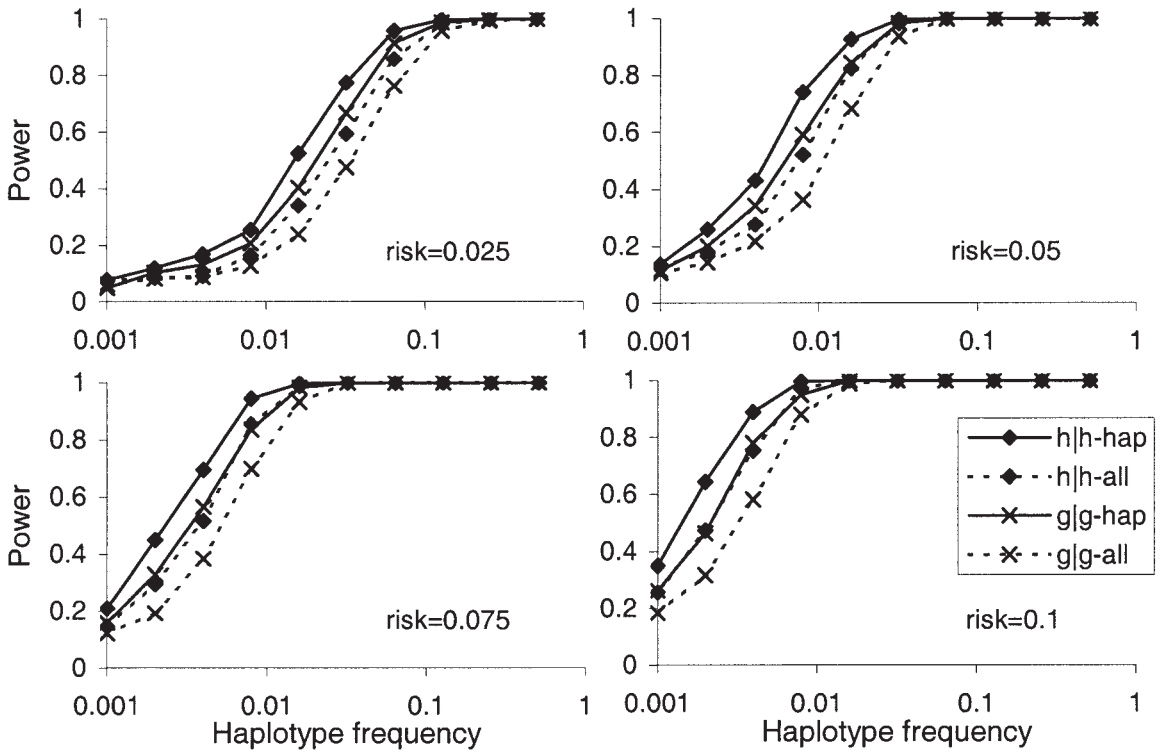

Fig. 2. Power to detect association using direct haplotyping or genotyping when simulated haplotype region consists of two biallelic marker loci. Different frequencies and different relative risks are shown for haplotype of interest. Solid lines (hap), power to detect association using a haplotype-specific test. Dashed lines (all), power in naive case. Simulation conditions: number of simulations per data point, 1,000; case sample size, 1,000; control sample size, 1,000; α=0.05.

4). This is due to two fundamental differences in the tests. Firstly, the specific test is dependent on the estimated standard error of the haplotype frequencies, while the naive test is not dependent on any standard errors. When the frequency of the haplotype is very low in the population and there are larger numbers of haplotypes (when $m>4$, for example), the standard errors become comparatively larger compared to h_d, thus decreasing the chance of finding a significant result. Secondly, the naive test examines the difference in distribution of the two sets of haplotype estimates, so that either very large differences in any of the haplotypes or small differences in a number of haplotypes can cause a significant result. At lower simulated values of h_d (e.g., <0.01), this results in a larger number of type 1 errors with respect to differences in h_d (i.e., there may be no significant difference in h_d, but fluctuations in other rare haplotypes may cause a significant result). At larger simulated values of h_d, there are larger differences in the frequency of d between groups, and so the relative number of type 1 errors falls.

NUMBER OF MARKER LOCI

For the two-locus case (Fig. 2), there is little difference in power to detect an association, with an approximately 10–20% increase in power when directly haplotyping for $h_d<0.1$, smaller than suggested by directly comparing relative efficiencies (see above). Clearly, at low haplotype frequency, d is very rare, so information on a statistical difference between case and control populations (or even difference from zero) is lacking. When $m>4$, the relative power to detect a difference using haplotyping is greatly increased, with relative haplotyping being seen to be three times as powerful with the specific test, and twice as powerful for the naive test at some values of h_d less than 0.1 when m=6 (Fig. 3). At these larger values for m, however, the absolute power to detect differences is reduced. In situations where the associated risk is about 0.075 or greater, haplotyping and genotyping show similar power to detect an association for $h_d>0.05$ (approximately), and therefore favor the use of genotyping. These trends continued when

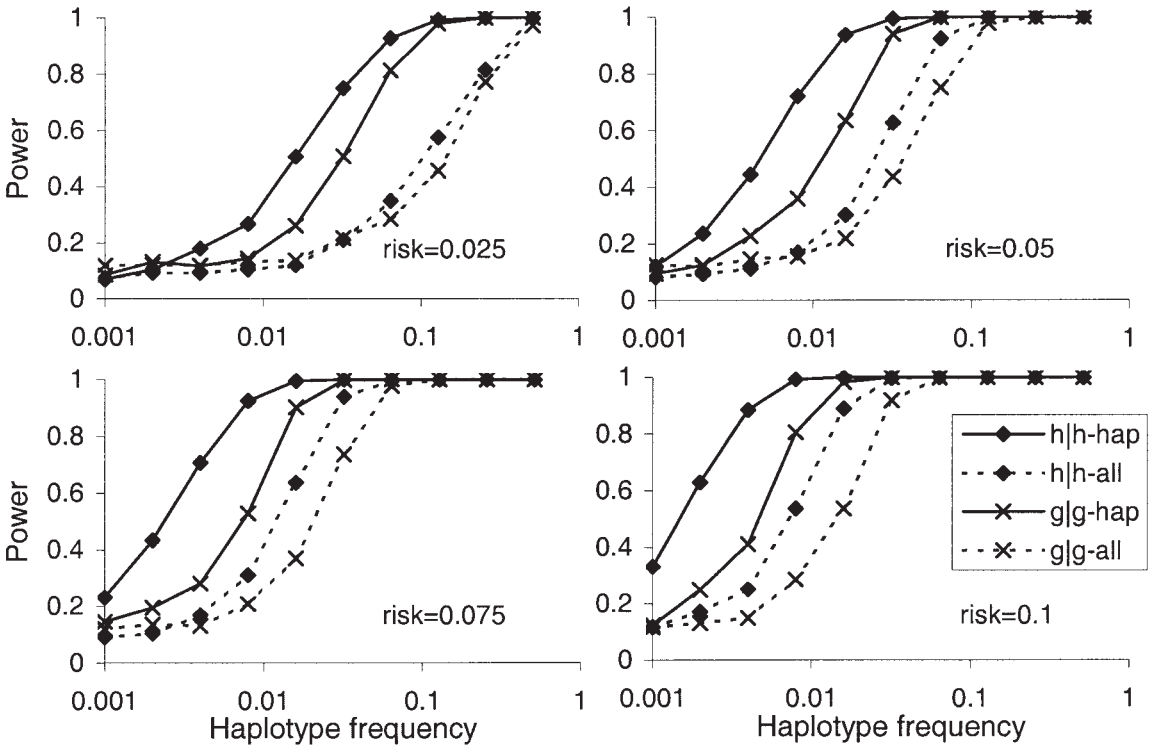

Fig. 3. Power to detect association using direct haplotyping or genotyping when region of interest contains six biallelic marker loci. Other simulation conditions as in Figure 2.

a smaller number of simulations was run for larger values for m (e.g., m=8; results not shown).

SAMPLE SIZE

Sample size influences the ability to detect associations (Fig. 4). At lower sample sizes (e.g., n=100 for both control and case groups), the actual power to detect associations is universally low, unless h_d is large (e.g., >10%) or R is larger. Unsurprisingly, absolute power increases when either sample size is increased, and particularly so when the lower of the two sample sizes is increased. For the naive test, incrementing either sample size has roughly the same effect on power. With the specific test, increasing the sample size of the case group yields a slightly larger increase in power, especially at lower values of h_d.

MIXING HAPLOTYPING AND GENOTYPING

Under either test for association, g | h and h | g lie between h | h and g | g (Fig. 5). They are roughly similar in magnitude, since sample size for case and control groups is identical, but with h | g being slightly greater overall. This was due to the slight advantage of accurately determining the frequency of the haplotype of interest in the case group. In the naive case, h | g and g | h lie closer to g | g than to h | h. This implies that haplotyping the case group while genotyping the control group does not provide a reliable means of increasing the relative power. If, however, the control group is much larger than the case group, then the relative power of h | g increases.

DISCUSSION

We investigated the relative power of direct haplotyping vs. genotyping to detect associations between a haplotype and a trait or disease locus under a range of simulated conditions, when there is linkage disequilibrium between markers in the sample. We showed that the relative power of haplotyping increases with decreasing sample size, decreasing linkage disequilibrium, and increasing numbers of marker loci. In addition, the frequency of the haplotype of interest and the magnitude of its association with the disease affects the power.

In the simulated studies presented here, direct haplotyping outperforms genotype analysis with

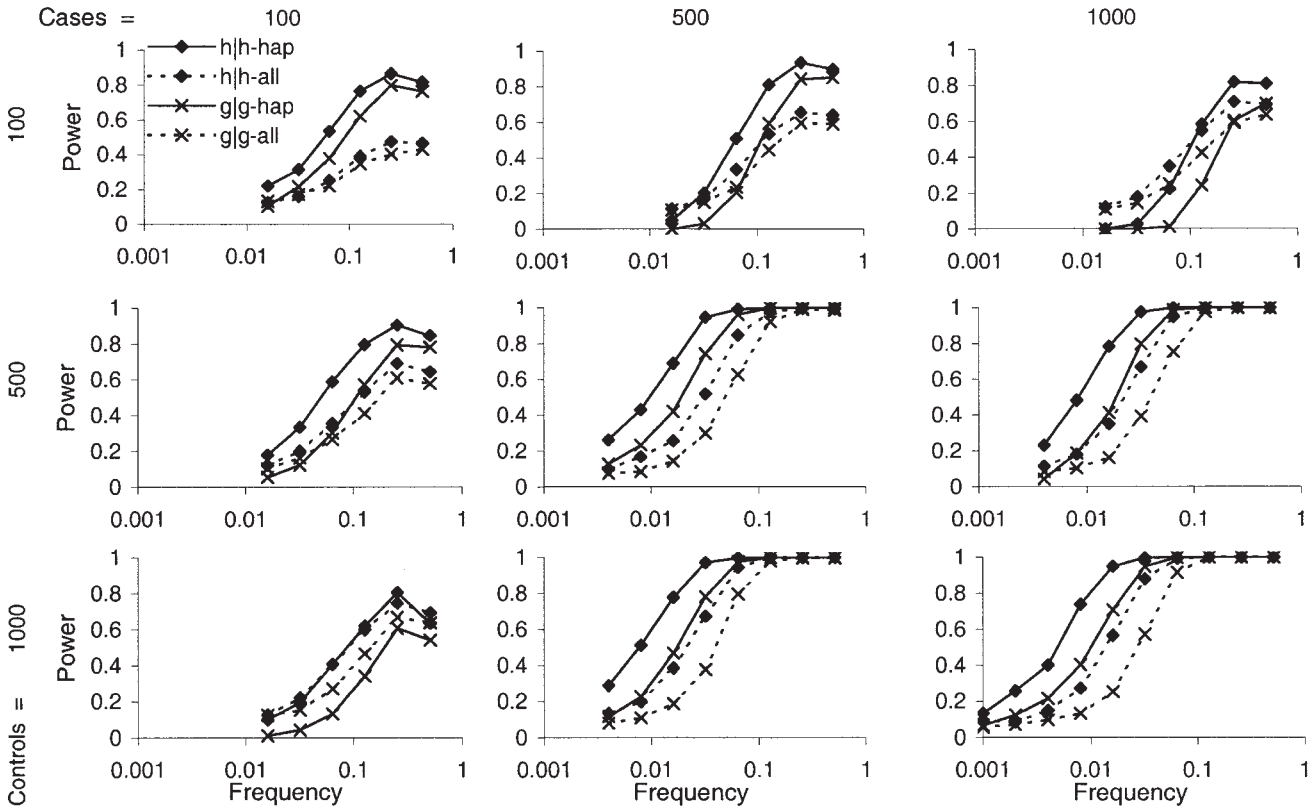

Fig. 4. Power to detect association using direct haplotyping or genotyping when case and control group sizes are varied. Range of h_d presented is where there is an average of at least two copies of d in both control and case groups. Simulation conditions: number of simulations per data point, 1,000; number of biallelic markers, 4; risk=0.05; α=0.05.

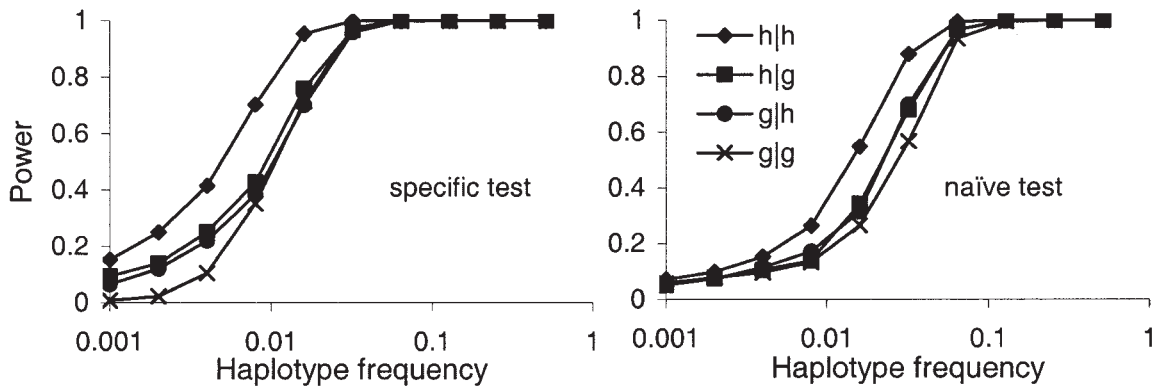

Fig. 5. Power to detect association using a combination of direct haplotyping and/or genotyping. Hap, power to detect when question of association with specific haplotype is addressed. All, power in naive case. Simulation conditions: number of simulations per panel, 10,000; case sample size, 1,000; control sample size, 1,000; number of biallelic markers, 4; risk=0.05; α=0.05.

respect to power per individual in the sample. However, at times when the relative power of haplotype analysis over genotyping is greatest, there is little absolute power by either method to detect subtle associations. Arguably the areas of most interest to the question of resource allocation lie in the steepest regions of the curves. These regions represent situations where the haplotype methodology outperforms genotyping, while still providing enough more absolute power to detect some more subtle associations. Above the steep regions there is little difference in power, and so genotyping would be favored. At other times, when the effect of d is larger (e.g., $R>0.075$),

genotyping will be almost as powerful as haplotyping to detect associations for $h_d > 0.05$. A circular problem arises in other situations, since the exact position of the steep region is unknown prior to study. Strong assumptions regarding the distribution of haplotypes, the size of h_d, and the size of R are required to resolve this problem. Empirical prior distributions for these parameters could be determined through meta-analysis of previous studies, provided there is a large selection of relevant previous studies. Realistically such an approach might lead to false assumptions, since there would be an upwards bias in the magnitude of associated effects detected previously (since only larger effects are reliably determined).

An additional problem for simulated studies on haplotypes is the need for a realistic method to simulate haplotype frequency. In addition to the Uniform approach described above, two further random distributions were investigated: Rectangular and Normal. These different distributions resulted in different numbers of alternate haplotypes being generated for each sample. The Rectangular and Uniform yielded a complete set of positive n_h, but Normal yielded a less than complete set (i.e., some $n_h=0$). In general, the greater the number of haplotypes simulated or the more regular their spread, the greater the relative power of haplotyping over genotyping, since under these conditions the percentage of ambiguous genotypes increases. Since the Uniform distribution always yields a sample containing at least one representative of each of the possible haplotypes, it is unlikely to realistically represent real-life situations. We argue, however, that the conclusion of low absolute advantage given uniform data can be extended to real-data situations. This is because efficiency values calculated using the Uniform distribution lay in the same region where haplotyping in the real-data situations showed highest relative efficiency. Ideally, prior information on the probability of observing a particular set of allele frequencies and their disequilibrium coefficients is required in order to weight simulations accordingly. Schork [2002] presented empirical distributions for disequilibrium coefficients in the two-locus case. How these generalize to the multilocus case when there are higher orders of disequilibrium coefficients is unknown.

Two obvious questions requiring further research arise. Firstly, what happens when money is more of a constraint than sample size? Is it then better to genotype as many individuals as possible, or haplotype fewer, or do a combination of both? Given the much larger laboratory resources required for direct haplotyping, genotyping would probably be favored, but this must be balanced against the substantial unit costs associated with recruitment and phenotyping. Moreover, noting the implicit prediction that paired haplotypes underpin etiology and outcome, interest in endophenotypes and subclassification of disease will grow in importance. Under such circumstances, the number of available cases would become the chief constraint rather than cost, and the power gained by conversion could be critical. Again in the context of multipurpose, prospective cohort studies, such as the planned UK Biobank study (www.wellcome.ac.uk/ukbiobank), there may be general value and gains from establishing a core set of haplotyped individuals to serve as controls for a number of case studies, or in using different case sets as controls for each other.

Secondly, what happens under more complex models of the disease? Different models for the disease will affect the relative power of the two tests examined. For example, the association of more than one haplotype with increased risk of a disease might improve the ability to detect associations using the naive test. More importantly, however, are the effects of the disease model on the relative efficiency of haplotyping. For example, the advantages of haplotyping over genotyping should increase in gene-gene interaction studies, since conversion methodologies provide a number of cell lines, each containing a random selection of single and paired human chromosomes.

The power study presented here is probabilistic in nature, and ultimately there are no guarantees that either haplotyping or genotyping will detect subtle genetic influences, even with large sample size and large numbers of marker loci. This caution is especially true in situations where the number of case individuals is limited simply because the trait or disease phenotype of interest is rare.

ACKNOWLEDGMENTS

This work was supported in part by the UK Medical Research Council (D.P. and P.M.V.) and the Cunningham Trust (D.P.). P.M.V. acknowledges support from the UK Biotechnology and

Biological Sciences Research Council. This project was undertaken by S.T. as a Special Study Module within the University of Edinburgh Undergraduate Course in Medicine. Real data were produced by the Chromosome 22 Group at the Sanger Institute, and were obtained from the World Wide Web at http://www.sanger.ac.uk/HGP/Chr22.

REFERENCES

Clark A. 1990. Inference of haplotypes from PCR-amplified samples of diploid populations. Mol Biol Evol 7:111–122.

Collins JE, Goward ME, Cole CG, Smink LJ, Huckle EJ, Knowles S, Bye JM, Beare DM, Dunham I. 2003. Reevaluating human gene annotation: a second-generation analysis of chromosome 22. Genome Res 13: 27–36.

Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL. 1993. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 261:921–923.

Dempster AP, Laird NM, Rubin DB,1977. Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc 39:1–38.

Douglas JA, Boehnke M, Gillanders E, Trent JM, Gruber SB. 2001. Experimentally-derived haplotypes substantially increase the efficiency of linkage disequilibrium studies. Nat Genet 28: 361–364.

Excoffier L, Slatkin M. 1995. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. Mol Biol Evol 12:921–927.

Fallin D, Schork NJ. 2000. Accuracy of haplotype frequency estimation for biallelic loci, via the expectation-maximization algorithm for unphased diploid genotype data. Am J Hum Genet 67:947–959.

Guo SW. 1997. Linkage disequilibrium measures for fine-scale mapping: a comparison. Hum Hered 47:310–314.

Hill WG. 1974. Estimation of linkage disequilibrium in randomly mating populations. Heredity 33:229–239.

Michalatos-Beloin S, Tishkoff SA, Bentley KL, Kidd KK, Ruano G. 1996. Molecular haplotypoing of genetic markers 10 kb apart by allele-specific long-range PCR. Nucleic Acids Res 24: 4841–4843.

Rannala B, Slatkin M. 1998. Likelihood analysis of disequilibrium mapping, and related problems. Am J Hum Genet 62:459–473.

Saunders AM, Schmader KE, Breitner JC, Benson MD, Brown WT, Goldfarb L, Goldgaber D, Manwaring MG, Szymanski MH, McCown N, Dole KC, Schmechel DE, Strittmayer WJ, Pericakvance MA, Roses AD. 1993. Apolipoprotein E 4 allele distributions in late-onset Alzheimer's disease and in other amyloid-forming diseases. Lancet 342:710–711.

Schaid DJ. 2002. Relative efficiency of ambiguous vs. directly measured haplotype frequencies. Genet Epidemiol 23: 426–443.

Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. 2002. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 70:425–434.

Schork NJ. 2002. Power calculations for genetic association studies using estimated probability distributions. Am J Hum Genet 70:1480–1489.

Sham P. 1998. Statistics in human genetics. Oxford: Oxford University Press, Inc.

Stephens M, Smith N, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989.

Yan H, Papadopoulos N, Marra G, Perrera C, Jiricny J, Boland CR, Lynch HT, Chadwick RT, De La Chapelle A, Berg K, Eshleman JR, Yuan W, Markowitz S, Laken SJ, Lengauer C, Kinzler KW, Vogelstein B. 2000. Conversion of diploidy to haploidy. Nature 403:723–724.