

Kevin Walters

The effect of genotyping error in sib-pair genomewide linkage scans depends crucially upon the method of analysis

Received: 26 April 2005 / Accepted: 7 June 2005 / Published online: 30 July 2005
© The Japan Society of Human Genetics and Springer-Verlag 2005

Abstract In genomewide linkage scans for complex diseases involving many loci with small genetic effects, it may be the case that no loci reach conventional statistical significance. A complementary method of evaluating linkage results, locus counting, may provide evidence for the existence of a number of genetic loci in these cases. Sib-pair study designs are often used in genomewide linkage scans, but because all genotype configurations are consistent with Mendelian inheritance, genotyping error will go largely undetected. Previous work on the effect of genotyping error has focused on a single disease locus. We considered the effect of two levels of genotyping error on genomewide evidence for linkage by using the simulated GAW 13 data. For affected sib-pair and non-parametric quantitative trait study designs, a 0.5% genotyping error rate reduced the number of independent linkage regions towards that expected under the null hypothesis of no linkage. A 2% genotyping error rate yielded less independent linkage regions than expected under the null hypothesis of no linkage. For a quantitative trait analysed using a parametric regression-based method, there was very little erosion of the linkage signal, even for error rates as high as 2%.

Keywords Genotyping error · Locus counting · QTL · ASP · Microsatellites · Simulation

Introduction

For complex diseases where multiple genes of modest effect size may be involved, many studies will have no

loci reaching conventional statistical significance (Altmüller et al. 2001). Evidence for multiple loci with relatively small genetic effects can be obtained using a locus-counting method (Wiltshire et al. 2002). This approach has been used in several studies (Caulfield et al. 2003; Frayling et al. 2003; Iyengar et al. 2004a, 2004b) to provide complementary evidence for linkage. In this approach, for a given LOD score, the number of independent regions of linkage (IRLs) expected under the null hypothesis of no linkage is determined by Monte Carlo simulation, and the number of independent linkage regions observed in the data is compared with this.

In all genetic studies, there will be some strategy employed to verify the accuracy of the genotyping. This may be achieved by checking that allele frequencies are in Hardy–Weinberg equilibrium (Hosking et al. 2004), comparing duplicated samples, independently calling alleles or by checking that the genotypes are consistent with Mendelian inheritance (Ewen et al. 2000). For diseases with a late age-at-onset, genomewide linkage studies utilising sibling pairs are a common study design due to the unavailability of parental DNA. Sib-pair study designs have the disadvantage that genotyping error is difficult to detect since all genotypic configurations are consistent with Mendelian inheritance. Methods for detecting unlikely genotypic configurations are available (Abecasis et al. 2002; Douglas et al. 2000), but in the absence of genotypes from other family members, many genotyping errors in sib-pair studies will go undetected. It is difficult to gauge the level of genotyping error in most genetic studies, as these figures are not routinely published. In a study to determine the levels and types of genotyping error typically encountered with microsatellite markers, Ewen et al. (2000) genotyped tens of thousands of microsatellite markers within families using both a commercial and custom-made marker set. By inheritance and concordance checking, they found genotyping error rates of 0.25% and 1.4% for the commercial and custom-made sets, respectively.

K. Walters
Mathematical Modelling and Genetic Epidemiology Group,
Division of Genomic Medicine, EU28,
School of Medicine & Biomedical Sciences, Beech Hill Road,
Sheffield, S10 2RX, UK
E-mail: k.walters@sheffield.ac.uk
Tel.: +44-114-2713046
Fax: +44-114-2711711

Clearly, these do not represent all the errors in the data, only those that are inconsistent with Mendelian inheritance or proved discordant when repeat samples were typed. The true error rates are almost certainly higher than the two figures quoted.

It has been demonstrated by simulation that genotyping error can substantially reduce LOD scores, particularly for affected sib-pair (ASP) designs (Abecasis et al. 2001a; Sullivan et al. 2003). Studies undertaken so far have simulated data for a single disease-linked locus and investigated the effect of genotyping error on the LOD score around this locus. We are interested in the effect of genotyping error on the number of linkage regions across the genome. Is the effect of a given level of genotyping error the same in genomewide terms as it is on a single locus? The locus-counting approach is becoming more widely used, and it is important to know the effect of genotyping error on this method. Could an inconclusive result be due to levels of genotyping error in the range typically encountered for microsatellite markers? We examine the effect of genotyping error on the number of independent linkage regions across the genome. The data are simulated under a model incorporating multiple disease loci of varying effect size under conditions typically encountered in an initial genomewide linkage scan using microsatellite markers. We consider study designs involving both quantitative trait loci (QTL) and ASP using multi-point calculations of IBD probabilities. We also examine a two-point analysis method for the ASP design.

Materials and Methods

Simulated data selection

The GAW 13 simulated data (Daw et al. 2003) mirror data encountered in an initial genomewide linkage scan, 10-cM map, 399 polymorphic microsatellite markers. The data were simulated without genotyping error. We will refer to these data as the true data. The phenotype we chose to analyse was the high-density lipoprotein (HDL) trait, as it had 11 loci affecting the trait phenotype with a range of genetic effect sizes. The 11 loci were distributed over seven chromosomes: three loci on chromosome 17, two on chromosomes 1 and 9 and one on chromosomes 3, 5, 11 and 21. For the quantitative trait analysis, we selected 4,700 sib-pairs from the replicates. We first adjusted for gender, smoking and drinking status (covariates known to influence the simulated HDL value) and used the residuals to represent the quantitative trait values. For the qualitative trait, we selected only those sib-pairs where both exceeded some threshold, chosen so as to provide a large enough sample of 728 sib-pairs. The figure of 4,700 sib-pairs for the quantitative trait was arbitrarily chosen so as to provide enough sib-pairs for the qualitative trait analysis.

Linkage analysis methods

All linkage analyses were undertaken using MERLIN (Abecasis et al. 2002). For the qualitative data, we performed linkage analysis using the method of Kong and Cox (1997) with a linear model. For the quantitative trait, we considered both a parametric and non-parametric method of analysis. The parametric method is the regression-based method of Sham et al. (2002). This method requires estimates of the population mean, variance and heritability. The sample values were used in each case (sample heritability was estimated from a variance components analysis of the true data). We choose the method of Sham et al. (2002) rather than variance components (VC) analysis because of the considerable computational time required to perform VC analysis on this many sib-pairs for 1,000 simulations (estimated at over 80 days). We could find no other work examining the effect of error on the method of Sham et al. (2002) whereas the effect on a VC analysis has been considered (Abecasis et al. 2001a), albeit under a single-locus disease model. To allow a crude comparison, we analysed 20 simulations using a VC method. For the qualitative and the two quantitative trait analysis methods, we undertook multi-point analysis. We also performed two-point analysis for the qualitative trait.

Locus-counting method

For each analysis method, we counted the number of independent regions showing linkage in the true data across an appropriate range of LOD scores. Regions were considered to be independent if their local maxima were more than 40 cM apart (the same criterion as Wiltshire et al. 2002). We then simulated 1,000 replicates of the data by the gene-dropping method (implemented in MERLIN, Abecasis et al. 2002) using the same markers, genetic map distances and allele frequencies under the null hypothesis of no linkage to any marker. For each LOD score in the range we considered and for each replicate, we counted the number of independent genomic regions with a local maximum above the given LOD score to obtain the empirical IRL distribution for each LOD score. To assess statistical significance, we could have used the 95th percentile of this empirical IRL distribution for each LOD score. However, it has been shown that for a given LOD score, the number of independent linkage regions under the null hypothesis of no linkage can be approximated by a Poisson distribution (Lander and Kruglyak 1995). Wiltshire et al. (2002) validated the use of the Poisson approximation in the sib-pair genome scan data they studied. Therefore, to assess statistical significance for each LOD score, we used the 95th percentile of the Poisson distribution with mean given by the mean of the 1,000 replicates. We chose to use the 95th percentile of the Poisson distribution for ease of calculation, as we needed the mean number of IRLs for each LOD score elsewhere.

In the two-point case, we counted the number of loci above a given LOD score and did not take proximity to other loci exceeding the threshold into account. Because the Poisson approximation in this case may not be appropriate, we used the 95th percentile of the empirical distribution of the number of markers exceeding the given LOD score as a measure of statistical significance. This guarantees a valid measure of statistical significance whether or not the Poisson approximation is valid in this case.

Incorporating genotyping error

For each analysis, we considered two rates of allelic error: 0.25% and 1%. These correspond to genotyping error rates of approximately 0.5% and 2%, respectively. The lower figure might represent an average genotyping error rate and the higher figure an upper boundary on the unknown error rate (based on the figures of Ewen et al. 2000). To allow comparison with other work in this area, we refer to the genotyping error rate rather than the allelic error rate in the rest of this paper.

There are several ways that genotyping error could have been incorporated into the GAW data. We wanted to use a simple model that would be general enough to allow for some of the most common types of error. Ewen et al. (2000) reported the type of microsatellite genotyping errors found after concordance and inheritance checking. They classified the genotyping errors into five categories: microsatellite mutations, priming site mutations, missed alleles, call errors and sample swaps. Sample swaps would lead to either one or more, probably two, of the alleles being misread. In genome-wide sib-pair linkage scans, many of these will be detected by relationship checking software such as GRR (Abecasis et al. 2001b) and RELATIVE (Goring and Ott 1997). Microsatellite mutations would generally lead to a unit repeat change in one of the alleles. Call errors (further classified as binning, leaking, low fluorescence, etc.) could lead to different types of genotyping error (homozygote to a heterozygote, heterozygote to a different heterozygote, etc.). The missed alleles and priming site mutations would generally lead to heterozygote genotypes being labelled as homozygote genotypes. We could have allowed each of these different types of error in our error model, but it is not clear what probabilities to assign to these different types of error. Because of the lack of studies focusing on this issue, we kept our model simple by allowing single allele changes rather than specifying probabilities for whole genotype changes. Our model allowed each allele to be misread with a fixed probability. If an allele was incorrectly read, then it was declared as one of the observed alleles at that marker with equal probability. One could change to another allele with probability equal to the allele frequency, but there is no evidence that this model reflects the true situation any more than the model we

used. If an allele is misread, why is it more likely to be read as one of the more common alleles? Our model allows for the possibility of a heterozygote being misread as a homozygote for one of the alleles and also allows (albeit with low probability) for heterozygote genotypes to be misread as heterozygote genotypes for different alleles (to represent undetected sample swaps). A drawback of our error model is the under-representation of missed allele and priming-site mutation type errors (where heterozygote genotypes may be misread as homozygote genotypes) according to the rates estimated by Ewen et al. (2000).

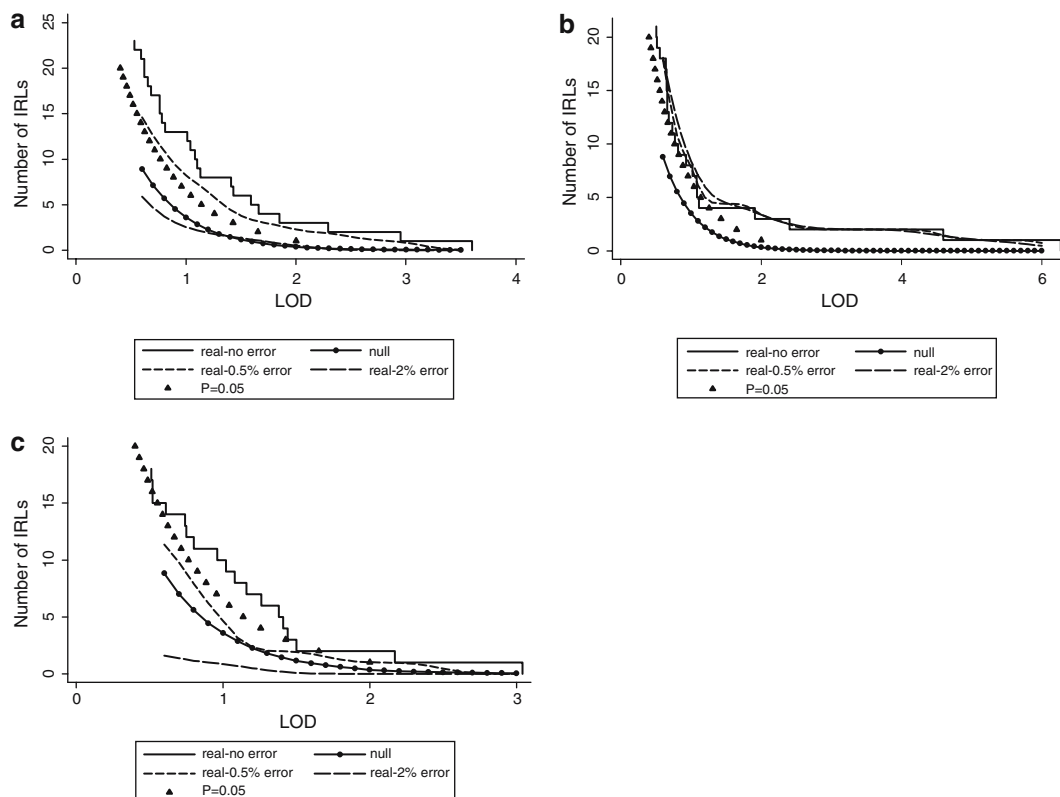
For each of the two genotyping error rates, we generated 1,000 replicates of the true data in MERLIN (Abecasis et al. 2002) and incorporated genotyping error into each simulation, as described. For each error rate and for a set of LOD scores covering the range of those observed, we counted the number of regions showing linkage for each simulation and averaged over the 1,000 replicates. In a true genome scan, allele frequencies ought to be calculated taking account of the familial relationships. Introducing genotyping error changes the allele frequencies, and so we would need to estimate the allele frequencies taking account of the familial relationship for each of the 1,000 replicates separately. This is not a realistic proposition in a genome scan comprised of 4,700 sib-pairs. We estimated allele frequencies for the replicates with genotyping error by using all individuals for each replicate separately. In order to make a fair comparison, we used the same method of allele frequency estimation with the true data.

We obtained the mean number of IRLs (and consequently the upper 5% tail of the Poisson distribution) by simulation using the allele frequencies from the true data. In a study that contained genotyping error, one would observe the data with error and simulate the null distribution based on the allele frequencies from these observed data. We simulated the null distribution based on allele frequencies from the true data containing no errors. The allele frequencies in the true data and the data with genotyping error will be different. Because of the way that genotyping error was simulated, it ought not to be clustered at particular markers, and because the rates are relatively low, the differences in allele frequencies should not be pronounced. We found that minor changes in the allele frequency distribution had very little effect upon the distribution under the null hypothesis. It is therefore reasonable to compare the results from the replicates with genotyping error with the distribution under the null hypothesis simulated using allele frequencies from the true data. As previously discussed, another potential genotyping error model involves changing a misread allele to another allele with probability equal to the allele frequency. One advantage of this alternative genotyping error model is that the expected allele frequencies of the data with simulated genotyping error would be the same as the observed data, thus obviating this problem.

Results

The effect of genotyping error on the actual number of IRLs for the QTL non-parametric, QTL regression-based and ASP cases analysed using multi-point IBD calculations are shown in Fig. 1. In all three cases, the number of independent linkage regions observed in the true data generally exceeds the 95th percentile of the appropriate Poisson distribution. Thus, in each case, it would be concluded that there is significant evidence for the presence of genetic loci contributing to the disease phenotype. It is difficult to compare the three methods directly, as there is varying evidence across the LOD score range in each case. To allow comparison across the different analysis types, the percentage of the number of IRLs retained is shown in Fig. 2. At both error rates, the

Fig. 1 **a** Number of independent regions of linkage (IRLs) for non-parametric quantitative trait loci (QTL) analysis: *solid line* corresponds to no error, *solid line with circle* to the expected number under the null hypothesis of no linkage, *short dashed line* to 0.5% error, *long dashed line* to 2% error and *solid triangles* to the 5% tail value of the appropriate Poisson distribution. **b** Number of IRLs for regression-based QTL analysis: *solid line* corresponds to no error, *solid line with circle* to the expected number under the null hypothesis of no linkage, *short dashed line* to 0.5% error, *long dashed line* to 2% error and *solid triangles* to the 5% tail value of the appropriate Poisson distribution. **c** Number of IRLs for qualitative trait multi-point analysis: *solid line* corresponds to no error, *solid line with circle* to the expected number under the null hypothesis of no linkage, *short dashed line* to 0.5% error, *long dashed line* to 2% error and *solid triangles* to the 5% tail value of the appropriate Poisson distribution



effect of error in the ASP case is more severe than in both of the QTL cases.

Non-parametric QTL

For the non-parametric QTL method, the variation in the percentage of the number of IRLs retained is much less than in the ASP case. Apart from the very top end of the LOD score spectrum, the effect of 0.5% genotyping error is to reduce the number of IRLs by approximately 20–40%. With this level of genotyping error, the number of IRLs is still well above the 95th percentile of the null IRL distribution. At the higher rate of error, the number of IRLs is generally less than the expected number under the null hypothesis, removing all evidence of genetic loci affecting the trait value. At this level of error, the number of IRLs is, on average, reduced to 75% of the number without error.

Regression-based QTL

For the regression-based method, the effect of error on the percentage of IRLs retained is more complicated. For some LOD scores, there is an increase in the number of IRLs. With 0.5% genotyping error rate, the percentage decrease in the number of IRLs is generally less than 20% and the increase generally less than 10%. The effect of the larger error rate on the number of IRLs is generally in the same direction as the lower error rate but is more pronounced.

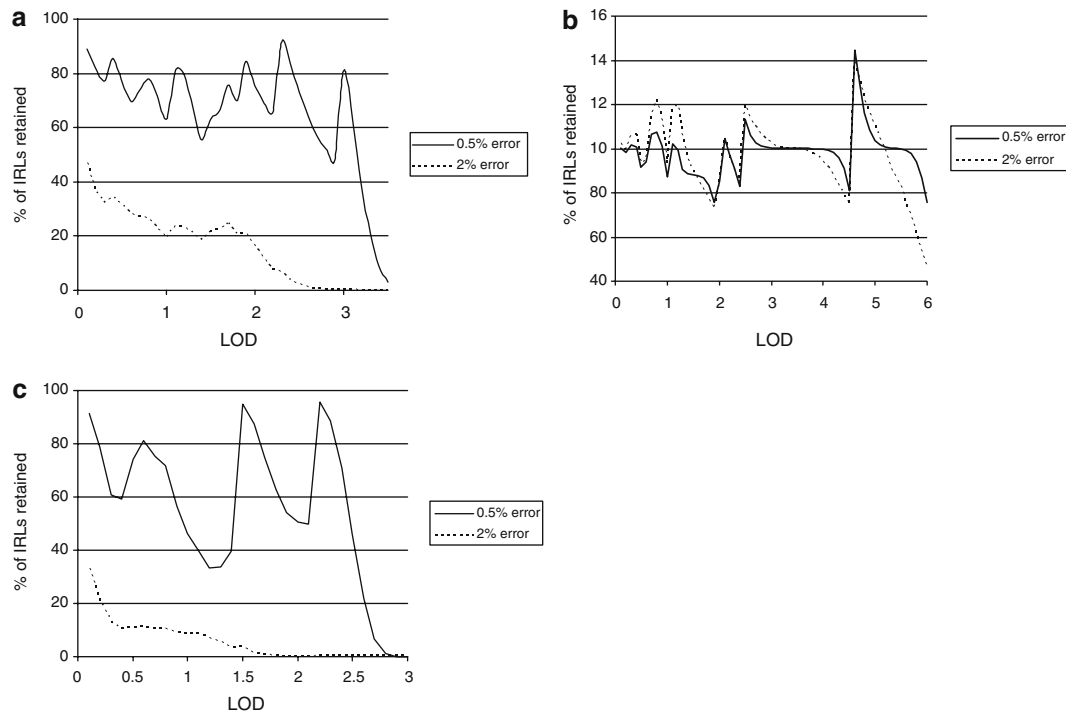


Fig. 2 **a** Percentage of independent regions of linkage (IRLs) retained for non-parametric quantitative trait loci (QTL) analysis: *solid line* corresponds to 0.5% error and *short dashed line* to 2% error. **b** Percentage of IRLs retained for regression-based QTL analysis: *solid line* corresponds to 0.5% error and *short dashed line* to 2% error. **c** Percentage of IRLs retained for qualitative trait multi-point analysis: *solid line* corresponds to 0.5% error and *short dashed line* to 2% error

Qualitative trait

For the ASP method, even at the smaller level of error, the number of IRLs decreases by more than 40% for about a half of the LOD score domain. This low rate of error brings the number of IRLs below the 95th percentile everywhere except for LOD scores exceeding 2 although the number of regions is still more than would be expected under the null hypothesis of no linkage across the whole LOD score range. At the higher error rate, the reduction in the number of linkage regions is more than 80% virtually everywhere, with a mean retention of around 10%. This higher rate of error removes all evidence of linkage such that the number of IRLs observed is considerably less than that expected under the null hypothesis for all LOD score values.

Comparing different QTL results

We looked at why the regression-based QTL method leads to increased mean numbers of IRLs at certain LOD scores but the non-parametric QTL method does not. To compare with the VC case, we also analysed 20 replicates of the data with a 0.5% genotyping error rate

using the VC method implemented in MERLIN (Abecasis et al. 2002). We determined some summary statistics for the effect of 0.5% genotyping error at two markers on one of the chromosomes. We chose one marker with a high and one with a low LOD score to see if the effect of genotyping error varied with the size of the LOD score. Table 1 shows the results for the high and low LOD scores. Given in Table 1 are the true LOD score, the mean and maximum LOD for the replicates with 0.5% error (percentage increase compared with the true LOD) and the number of replicates with a LOD score greater than the true LOD score. We checked that none of the replicates produced higher LOD scores than the true score in the ASP case.

For the three QTL methods, the results show that it is both the size of the increase and the probability that genotyping error causes an increase that explains why the regression-based method leads to a mean increase in the LOD score whilst neither the non-parametric method nor the VC method do. For QTLs at markers with low LOD scores, the presence of genotyping error will result in a greater chance of the error increasing the LOD score compared with genotyping error around markers yielding higher LOD scores. The size of the effect is also likely to be greater around markers producing low LOD scores. These findings are reinforced by Fig. 1b. For the regression-based method, compared with the case without error, genotyping error increases the number of linkage regions only at lower LOD scores. Due to the limited number of observations for the VC analysis, the true LOD score is less than one standard error away from the sample mean in both cases and so the results must be treated cautiously. However, this reduction in LOD score for the VC case in the presence

Table 1 The effect of 0.5% genotyping error at high and low LOD scores for a quantitative trait loci (QTL).^a VC variance components

	LOD	Mean	Maximum	Number exceeding true LOD
High LOD				
QTL non-parametric	3.60	3.15	3.73 (3.6%)	4/1,000
QTL regression-based	4.59	4.55	5.35 (16.6%)	412/1,000
VC	5.71	5.67	6.32 (10.7%)	8/20
Low LOD				
QTL non-parametric	0.670	0.507	0.74 (10.4%)	18/1,000
QTL regression-based	0.803	0.806	1.15 (43.2%)	491/1,000
VC	0.790	0.786	0.95 (20.2%)	9/20

^aLOD represents the score without error; the last three columns refer to the simulations with error; in brackets, the percentage increase over the true LOD

of genotyping error is consistent with the observations of Abecasis et al. (2001a).

Two-point analysis

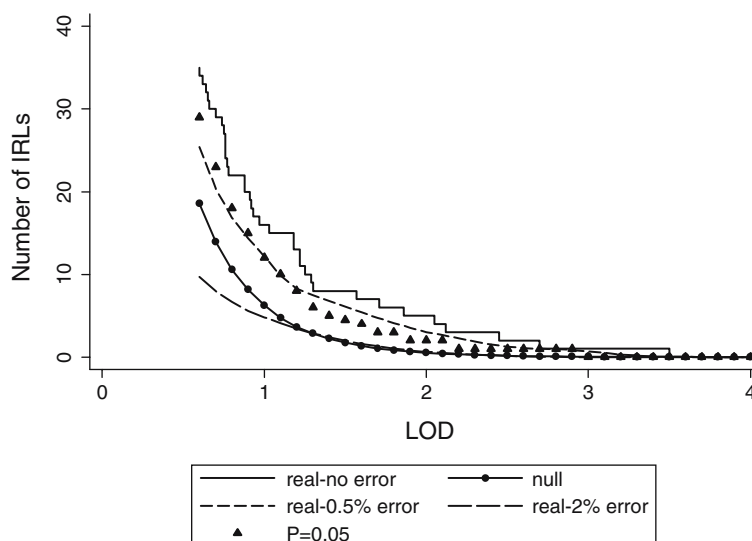
We also performed a two-point ASP analysis. IBD sharing probabilities were calculated at each marker using only the genotype data for that marker. Ignoring the inheritance information at surrounding markers has the disadvantage that information content may be considerably diminished but has the advantage of being unaffected by genotyping error should it exist in surrounding markers. Since the IBD probabilities were calculated using the marker data only, we simply counted the number of markers exceeding the LOD score and did not take the LOD scores of neighbouring markers into account. As a measure of statistical significance, we used the 95th percentile of the empirical distribution of the number of markers exceeding the given LOD. The results are shown in Fig. 3. As in the multi-point case with no error, the number of markers exceeding a given LOD score is in the 5% tail of the empirical distribution, providing evidence for genetic effects across the entire LOD score range considered. A

genotyping error rate of 0.5% does not generally reduce the number of markers below the 95th percentile for most LOD scores. This was not the case for the ASP multi-point analysis where this level of error brought the number of IRLs below the 95th percentile, removing much of the significant evidence for an excess in the number of independent linkage regions. The effect of 2% genotyping error is also less extreme in the two-point compared with the multi-point case. For this error rate, the number of markers with a LOD score above a given value is similar to that obtained by simulating under the null hypothesis of no linkage.

Effect of genotyping error on the position of the maximum LOD score

As well as looking at the effect of genotyping error on the strength of evidence for linkage, we also investigated the localisation of the true linkage peaks. The GAW data contained 11 trait loci distributed over seven chromosomes. All four (a two-point and three multi-point) types of analysis without genotyping error correctly identified either four or five true loci with a LOD score exceeding 1. Between three and nine false positives

Fig. 3 Number of independent regions of linkage (IRLs) for qualitative trait two-point analysis: *solid line* corresponds to no error, *solid line with circle* to the expected number under the null hypothesis of no linkage, *short dashed line* to 0.5% error, *long dashed line* to 2% error and *solid triangles* to the 5% upper tail value of the empirical IRL distribution



were identified. For the two-point analysis, the average marker spacing was 10 cM. For the multi-point analysis, LOD scores were also calculated at two equally spaced distances between all genotyped markers, as might be done in a true linkage scan, to allow for more accurate positioning of the maximum LOD score. Spacing between LOD score calculations was therefore 3.3 cM for the multi-point analysis.

In the multi-point case, for each of the correctly identified loci, we examined the LOD scores in a region spanning five positions (16.5 cM) either side for each replicate. Because of the greater spacing of points at which LOD score calculations are performed in the two-point analysis, we considered only two markers either side of the true marker. We calculated the percentage of replicates for which the position with the highest LOD score in this region was different to the position identified in the analysis without genotyping error. As well as examining how frequently the position of the maximum LOD score changed in this region, we were also interested in the distance between the true position and the suggested position for each replicate. We calculated the percentage of replicates (where the position had been incorrectly specified) in which the position of the maximum LOD score changed by more than one. For each method of analysis, we combined the results for all the correctly identified loci in the true data. The results are given in Table 2.

For brevity, Table 2 gives the mean across all true disease loci. There was considerable variation at individual loci that was not explainable either in terms of the size of the true LOD scores or in the level of genotyping error. It might be expected that the proportion of incorrectly specified positions and the size of the misspecification would increase with the level of genotyping error and that higher LOD scores would be more robust to genotyping error in terms of localisation; this was not consistently observed. Although there were no obvious patterns within the analysis methods, there are generalisations that can be made between the different methods. From Table 2 it is clear that at both error rates, the regression-based QTL analysis was most affected in terms of disease localisation. It had the highest proportion of replicates that incorrectly specified the disease location and the highest proportion that

incorrectly specified by more than one position. The multi-point analysis of the qualitative trait was the next worst performer in terms of localisation, particularly at the higher error rate. At the lower error rate, the positions of the highest LOD scores for both the non-parametric QTL and two-point analysis of the qualitative trait were largely unaffected by genotyping error. At most, 5% of the replicates incorrectly specified the position of the true locus and never by more than one position. The non-parametric QTL method is clearly the most robust of the multi-point methods in terms of disease localisation. Even with a 2% genotyping error rate, 85% of replicates correctly specify the location of the disease gene, and of those that do not, 86% are only 3.3 cM away. Although the proportion of incorrectly identified loci is quite low for the two-point analysis, the consequences are potentially severe because of the spacing of the markers; however, where a genome scan is undertaken, two-point analysis would usually be used in conjunction with a multi-point analysis of the trait.

Discussion

In our simulated genomewide linkage scan involving a quantitative trait, where the Sham et al. (2002) method implemented in MERLIN (Abecasis et al. 2002) was used and a locus-counting approach employed as a complementary method of assessing evidence for linkage, the presence of 0.5–2% genotyping errors had very little net effect on the number of IRLs across the whole LOD score range. Indeed, at low LOD scores, genotyping error was found to increase the mean number of IRLs. For the non-parametric quantitative trait analysis implemented in MERLIN (Abecasis et al. 2002) and the affected sib-pair design using the method of Kong and Cox (1997), the effects were considerably more severe.

Our results, combined with those of Abecasis et al. (2001a), show that genotyping error (percentage reduction) has a more severe effect upon the number of linkage regions than on the actual LOD score at a disease-linked marker. From our own results (not shown), we see that under an ASP study design, a locus providing evidence suggestive of linkage could remain suggestive of linkage in the presence of a genotyping error rate of 0.5%; a LOD score of 3.04 reduces to a mean LOD score of 2.48 under a 0.5% genotyping error rate, both of which are suggestive of linkage under the guidelines suggested by Lander and Kruglyak (1995). Using a locus-counting method, this level of error changes the interpretation considerably. Where with no error the locus-counting results provide complementary evidence for a genetic basis for the disease, with just 0.5% error there is virtually no evidence of a genetic component, as the number of independent linkage regions is little more than that expected under the null hypothesis of no linkage. So it is possible in ASP studies for the results from a linkage scan to provide several loci

Table 2 The percentage of replicates for which the marker with the greatest LOD score in a region surrounding a true locus is not the correct marker. *QTL* quantitative trait loci

	0.5% error rate (%) ^a	2% error rate (%) ^a
Non-parametric QTL	2 (0%)	15 (14%)
Regression-based QTL	33 (34%)	46 (40%)
Qualitative trait (multi-point)	6 (1%)	38 (34%)
Qualitative trait (two-point)	5 (0%)	13 (5%)

^aIn brackets, the percentage of those replicates incorrectly specifying the location in which the location is incorrectly specified by more than one marker is given

with results suggestive of linkage according to the conventional genomewide significance levels (Lander and Kruglyak 1995) but for the number of independent regions showing evidence of linkage not to significantly exceed the number expected under the null hypothesis of no linkage. Using an ASP design, a lack of evidence using the locus-counting method could be due to genotyping error at a rate of around 1%. The same comments apply to the non-parametric QTL analysis method although the genotyping error rate would probably have to exceed 1–1.5% in order for all evidence of genetic effects to be removed. Particularly for the ASP case, using a locus-counting method may provide evidence of genotyping error. If the number of independent regions showing evidence of linkage is less than the null across a substantial proportion of the LOD score range, then this may indicate the presence of genotyping error. Under an affected-sib-pair design using a two-point analysis, the reduction in the number of linkage regions is less than in the multi-point case, but error rates of 1–2% would still make the results difficult to interpret.

In terms of localisation of disease genes, the robustness of the regression-based method does not hold. It performed worse than all the other methods in terms of the proportion of replicates incorrectly specifying the disease gene location and in terms of the distance of the suggested location from the true location of the maximum LOD score. The consequences of this will depend to some extent on the follow-up from the initial linkage scan. Often, a further linkage analysis is performed using a more dense set of markers in regions of interest to further define a putative disease-linked locus. If these new markers span a large enough distance (15 cM, say), then the true position of the locus may be included within this set.

The ASP design is more severely affected by error, as error tends to reduce the mean IBD sharing. This reduces the evidence for linkage. For the quantitative trait methods, the presence of genotyping error can increase the linkage signal and so the mean effect on the number of IRLs is less severe. The implications of this are that in the presence of genotyping error, some false positives may be inflated under QTL designs and some true positive LOD scores may be diminished. In the case of QTLs, the locus-counting method has the advantage that counting the number of regions across the genome will act to balance these fluctuations in the number of independent linkage regions. This may give a more accurate genomewide result when there are locus-specific inflations and reductions in LOD scores due to genotyping error.

The results from a locus-counting approach can be difficult to interpret unless the number of IRLs exceeds the 95th percentile of the Poisson distribution with appropriate mean across a large proportion of the LOD score range. These results show that just 0.5% genotyping error rate can make the interpretation difficult, particularly under an ASP study design. These simulations assume that no genotyping error is found. With sib-pairs and no other relatives, all genotypic configura-

tions are consistent with Mendelian inheritance, and error detection relies upon looking for unlikely genotypes, checking for Hardy–Weinberg equilibrium, etc. With other family members, some erroneous genotypes will be detected and either corrected or removed from the analysis resulting in lower genotyping error rates than those considered in this study.

The conclusions of this paper are of course limited to the particular genetic model used in the simulation of the GAW 13 data and therefore cannot necessarily be generalised to other models. This will always be the case since in order to examine the effect of genotyping error, one has to start with clean (i.e. simulated) data. However, the model used to simulate the GAW 13 data represents a considerably more sophisticated multi-locus model than the single-locus cases that have so far been considered in the context of estimating the effect of genotyping error. The results also depend upon the genotyping error model used. The allele-change model we use under estimates certain types of error and over estimates others, as will all error models. We could have based our model on, for example, the work of Ewen et al. (2000), but then the conclusions would not apply to studies where genotyping error manifests itself in a different form. The aim of this paper was to compare the effects of genotyping error on the genomewide number of linkage regions using different methods of analysis rather than to examine the effect of different error models.

The effect of genotyping error depends upon the distribution of errors as well as the rate. In this study, genotyping error is approximately uniformly distributed over all markers. In reality, the error may be more clustered at “poor” markers. The effect of this error clustering will depend upon the proximity of the cluster to a true disease susceptibility locus. This paper emphasises the need to minimise genotyping error using all available techniques, as no methods yet exist for the inclusion of genotyping error in model-free linkage analysis. Genomewide linkage scans using SNP markers are now a feasible study design (John et al. 2004). Given sufficient density, SNP markers can extract as much of the inheritance information as a standard 10-cM microsatellite map (Evans and Cardon 2004) and are subject to much lower rates of genotyping error. In light of the results of this study, it may be that a genomewide scan using SNPs that extracts a lower percentage of the inheritance information than a 10-cM microsatellite scan may provide more complementary evidence for linkage, assessed by locus counting, due to the lower genotyping error rate associated with this technology although this needs further consideration.

References

- Abecasis GR, Cherny SS, Cardon LR (2001a) The impact of genotyping error on family-based analysis of quantitative traits. *Eur J Hum Genet* 9:130–134

- Abecasis GR, Cherny SS, Cookson WOC, Cardon LR (2001b) GRR: graphical representation of relationship errors. *Bioinformatics* 17:742–743
- Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30:97–101
- Altmüller J, Palmer LJ, Fischer G, Scherb H, Wjst W (2001) Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 5:936–950
- Caulfield M, Munroe P, Pembroke J, Samani N, Dominiczak A, Brown M, Benjamin N, Webster J, Ratcliffe P, O’Shea S, Papp J, Taylor E, Dobson R, Knight J, Newhouse S, Hooper J, Lee W, Brain N, Clayton D, Lathrop GM, Farrall M, Connell J (2003) Genomewide mapping of human loci for essential hypertension. *Lancet* 361:2118–2123
- Daw EW, Morrison J, Zhou X, Thomas DC (2003) Genetic analysis Workshop 13: simulated longitudinal data on families for a system of oligogenic traits. *BMC Genet* 4(Suppl 1):S3
- Douglas JA, Boehnke M, Lange K (2000) A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. *Am J Hum Genet* 66:1287–1297
- Evans DM, Cardon LR (2004) Guidelines for genotyping in genomewide linkage studies: single-nucleotide-polymorphism maps versus microsatellite maps. *Am J Hum Genet* 75:687–692
- Ewen KR, Bahlo M, Treloar SA, Levinson DF, Mowry B, Barlow JW, Foote SJ (2000) Identification and analysis of error types in high-throughput genotyping. *Am J Hum Genet* 67:727–736
- Frayling TM, Wiltshire S, Hitman GA, Walker M, Levy JC, Sampson M, Groves CJ, Menzel S, McCarthy MI, Hattersley AT (2003) Young-onset type 2 diabetes families are the major contributors to genetic loci in the Diabetes UK Warren 2 genome scan and identify putative novel loci on chromosomes 8q21, 21q22, and 22q11. *Diabetes* 52:1857–1863
- Goring HHH, Ott J (1997) Relationship estimation in affected sib-pair analysis of late-onset diseases. *Eur J Hum Genet* 5:69–77
- Hosking L, Lumsden S, Lewis K, Yeo A, McCarthy L, Bansal A, Riley J, Purvis I, Xu CF (2004) Detection of genotyping errors by Hardy–Weinberg equilibrium testing. *Eur J Hum Genet* 12:395–399
- Iyengar SK, Song DH, Klein BEK, Klein R, Schick JH, Humphrey J, Millard C, Liptak R, Russo K, Jun G, Lee KE, Fijal B, Elston RC (2004a) Dissection of genomewide-scan data in extended families reveals a major locus and oligogenic susceptibility for age-related macular degeneration. *Am J Hum Genet* 74:20–39
- Iyengar SK, Klein BEK, Klein R, Jun G, Schick JH, Millard C, Liptak R, Russo K, Lee KE, Elston RC (2004b) Identification of a major locus for age-related cortical cataract on chromosome 6p12–q12 in the Beaver Dam Eye Study. *Proc Natl Acad Sci USA* 101:14485–14490
- John S, Shephard N, Liu GY, Zeggini E, Cao MQ, Chen WW, Vasavda N, Mills T, Barton A, Hinks A, Eyre S, Jones KW, Ollier W, Silman A, Gibson N, Worthington J, Kennedy GC (2004) Whole-genome scan, in a complex disease, using 11,245 single-nucleotide polymorphisms: comparison with microsatellites. *Am J Hum Genet* 75:54–64
- Kong A, Cox NJ (1997) Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 61:1179–1188
- Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11:241–247
- Sham PC, Purcell S, Cherny SS, Abecasis GR (2002) Powerful regression-based quantitative-trait linkage analysis of general pedigrees. *Am J Hum Genet* 71:238–253
- Sullivan PF, Neale BM, Neale MC, van den Oord E, Kendler KS (2003) Multipoint and single point non-parametric linkage analysis with imperfect data. *Am J Med Genet* 121B:89–94
- Wiltshire S, Cardon LR, McCarthy MI (2002) Evaluating the results of genomewide linkage scans of complex traits by locus counting. *Am J Hum Genet* 71:1175–1182