Program Update and Novel Use of the DESPAIR Program to Design a Genome-Wide Linkage Study Using Relative Pairs

Heather M. Ochs-Balcom^{a, b} Xiuqing Guo^c Takashi Yonebayashi^c
Georgia Wiesner^d Robert C. Elston^{b, e}

^aDepartment of Social and Preventive Medicine, School of Public Health and Health Professions, University at Buffalo, Buffalo, N.Y., ^bDepartment of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, Ohio, ^cMedical Genetics Institute, Burns and Allen Research Institute, Cedars-Sinai Medical Center, Los Angeles, Calif., ^dCenter for Human Genetics and Department of Genetics, University Hospitals Case Medical Center and Case Western Reserve University, and ^eIreland Comprehensive Cancer Center at University Hospitals Case Medical Center, Case Western Reserve University, Cleveland, Ohio, USA

Key Words

Linkage analysis • Power • Sib pair linkage analysis • Model-free linkage analysis • DESPAIR • Relative pairs

Abstract

DESPAIR (DESIGN for PAIRs) is a computer program useful for designing a two-stage linkage study using relative pairs for a dichotomous phenotype. It determines the optimal two-stage study design – i.e., for specified power and significance level, how many pairs of relatives should be studied, how many equally spaced markers should be used initially, and what criterion should be used to specify the markers around which further searching should be done at a second stage. The program will calculate either the number of relative pairs required for a given number of first-stage markers or the number of markers required for a given number of relative pairs. We highlight the use of the latest version of DESPAIR to decide to what extent additional fine mapping in a candidate region of interest can lead to an increase in power in a previous linkage study sample. We also discuss new features of the program, such as the mean difference test for affected and discordant relative pairs and estimation

of full sibling pair equivalents to design a study when several types of relative pairs are available. DESPAIR is part of the S.A.G.E. program package and is freely available for use online.

Copyright © 2009 S. Karger AG, Basel

DESPAIR (DESIGN for PAIRs) is a program for designing a model-free linkage study for a dichotomous phenotype using concordantly affected (originally) or both concordant and discordant (now) relative pairs. Users input relative recurrence risks of interest, the relative cost of phenotyping and genotyping individuals, desired power and significance level, and other parameters to estimate the number of relative pairs that should be studied, how many equally spaced markers should be used, and what criterion should be used to specify the markers around which further searching should be done at a second stage. Alternatively, for specified power and significance level, the program will calculate either the number of pairs required for a given number of first-stage markers or the number of markers required for a given number of pairs.

We summarize the new features and utility of DESPAIR and highlight a novel use for making decisions about whether fine mapping is advisable in particular genomic regions of interest, such as those identified by prior linkage studies or candidate genes that are biologically relevant to disease risk. Users can now utilize tests based on mean identical by descent (IBD) allele sharing as well as the proportion of relative pairs sharing 0 alleles IBD, and can incorporate discordant pairs into their study design. We show how non-full sibling pairs can be converted into approximate sibling pair equivalents, which may be especially useful in samples where the available relative pair types are varied. We discuss two different ways to allow for heterogeneity, current DESPAIR limitations and plans for future program development.

The new DESPAIR program is freely available for online use by clicking on the DESPAIR link on the S.A.G.E. website <http://darwin.cwru.edu/despair/>. To begin, the user must click on the 'Create Design' link and input the number of scenarios to be considered, which will facilitate the comparison of resulting output utilizing different assumed parameters. The user must specify the desired significance level α , the desired power, whether an approximate or exact method is to be used to calculate power, the type of sample of interest, including the type of relative pair (full-sib, half-sib, grandparent-grandchild, first cousin or avuncular) and the concordance type (concordantly affected relative pairs, discordant relative pairs, or both). The original DESPAIR program was designed for samples of affected relative pairs only [1], but is now extended to have part or the entire sample made up of discordant relative pairs [2].

The user must specify the *locus-specific* relative recurrence risk ratios for the offspring and siblings of an affected person, λ_O and λ_S , respectively, to indicate the magnitude of the effect to be detected at any particular locus. These are defined as the probability that an offspring or sibling of an affected person is affected divided by the population prevalence of the disease. Originally specified for a parent or offspring, this relative recurrence risk can now also be specified for a sibling.

Originally based on the proportion of relative pairs sharing 0 alleles IBD, DESPAIR is now extended to include the option to base a linkage test on the mean proportion of alleles shared IBD which, in the case of full sibling pairs leads to a more powerful test when recurrence risk to first degree relatives is small [3, 4].

Assume that m equally spaced markers are typed at the first stage on $2n$ individuals (n pairs) and that $2k$ ad-

ditional markers are typed around each first-stage marker that demonstrates a p value less than a calculated α^* ; thus, on each side of such a marker, the distance between the marker and halfway to the next marker is divided into $k + 1$ equal intervals by the k additional markers. Then the total expected study cost in units of genotyping one marker on one individual was taken to be proportional to

$$C = 2n\{R + m + 2k[\alpha^* \cdot m + (1 - \beta)d]\}, \quad (1)$$

where R is the ratio of the cost of recruiting (and phenotyping) an individual to that of genotyping one marker on one individual, $1 - \beta$ is the desired power and d is the number of disease loci present in the whole genome. The power is to detect a specified relative recurrence risk, due to segregation at the linked disease locus.

In the original version of DESPAIR, the α^* that minimizes cost function (1) for various values of k assumes conservatively that a particular disease locus lies up to halfway towards the next first-stage marker on either side. Ziegler et al. [5] pointed out that the cost function (1) ignores the fact that there may be more than one first-stage marker linked to the disease locus at any one point, and so the total expected cost is more appropriately reflected by

$$C = 2n\left\{R + m + 2k\left[\alpha^*\left(m - \sum_{i=1}^d l_i\right) + \sum_{i=1}^d \sum_{j=1}^{l_i} (1 - \beta_{ij})\right]\right\}, \quad (2)$$

where l_i is the number of first-stage markers linked to disease locus i , and $1 - \beta_{ij}$ is the probability that $2k$ second-stage markers are typed around marker j that is linked to disease locus i at significance level α^* . Users now have the option to input a maximum distance (g) between any disease locus and a 'linked' marker. Then significant results obtained within g centimorgans from any disease locus are considered to be successes and any outside that range are considered to be false positives. By making the distance g small in comparison to the distance between first-stage markers, cost function (2) approaches cost function (1), which was the function used in the original DESPAIR.

Assume that no two disease loci are closer than $2g$ centimorgans apart and conservatively that the disease locus lies halfway between two adjacent first-stage markers. Then the number of linked markers within this distance is

$$l = 2\left\{\left\lceil \frac{g - x}{2x} \right\rceil + 1\right\},$$

Fig. 1. Distance between the j -th marker (on either side) and the disease locus. The distance between the first marker ($j = 1$) and the disease locus is $x = 2^*(1-1)*x + x$; the distance between the second marker ($j = 2$) and the disease locus is $3x = 2^*(2-1)*x + x$; and so on.

where $2x$ is the distance between two adjacent markers and $[(g - x)/(2x)]$ denotes the largest integer that is smaller than $(g - x)/(2x)$. For full sibling pairs, the deviation from the mean proportion of alleles shared IBD under the null hypothesis at the j -th marker on either side of the i -th disease locus is given by

$$\delta_{ij}(\theta_{ij}) = \delta_j(\theta_{ij}) = (2[\theta_{ij}^2 + (1 - \theta_{ij})^2] - 1)\delta, \quad (3)$$

where, in terms of the locus-specific relative recurrence risks λ_S and λ_O , respectively for sibs and offspring, $\delta = (2\lambda_S - \lambda_O - 1)/(4\lambda_S)$ [6]; and the recombination fraction θ_{ij} between the i -th disease locus and the j -th marker on either side of it is derived from the genetic distance between the two loci using a map function K , i.e. $\theta_{ij} = K(\text{distance})$. The distance between the disease locus and the j -th first stage marker on each side is $2(j - 1)x + x$ (see fig. 1). For all types of relative pairs, the deviation from the proportion of relative pairs who share 0 alleles IBD under the null hypothesis can be similarly calculated, with δ as given in Guo and Elston [3]. In either case, the probability of typing $2k$ markers at the second stage around the j -th marker that provides evidence for linkage with the i -th disease locus at the first stage is

$$1 - \beta_{ij} = 1 - \beta_j = 1 - \Phi \left[\frac{\frac{1}{2} Z_{\alpha^*} - \delta_j \sqrt{2n}}{\sqrt{\frac{1}{4} - \delta_j^2}} \right],$$

where Φ is the standard normal distribution function. Given values of all the other parameters, this equality implicitly defines α^* . The total expected cost can be simplified to

$$C = 2n \left\{ R + m + 2k \left[\alpha^* (m - dl) + 2d \sum_{j=1}^{l/2} (1 - \beta_j) \right] \right\}, \quad (4)$$

as explained by Ziegler et al. [5], and using this cost function the overall power to detect linkage is greater than $1 - \beta$, i.e., the optimal design obtained was conservative,

Table 1. Deviation from the null hypothesis mean proportion of alleles shared IBD at the first marker on either side of a disease locus when contrasting affected and discordant pairs; p is the proportion of relative pairs that are discordant

Type of relative pair	Δ
Full sib	$\frac{(2\psi - 1)}{4(p + 1)} \left[\frac{p(2\lambda_S - \lambda_O - 1)}{\lambda_S} - \frac{(2\lambda_S^- - \lambda_O^- - 1)}{\lambda_S^-} \right]$
Grandparent-grandchild	$\frac{(1 - 2\theta)}{4(p + 1)} \left[\frac{p(\lambda_O - 1)}{1 + \lambda_O} - \frac{(\lambda_O^- - 1)}{1 + \lambda_O^-} \right]$
Avuncular	$\frac{(1 - \theta)(1 - 2\theta)^2}{4(p + 1)} \left[\frac{p(\lambda_O - 1)}{1 + \lambda_O} - \frac{(\lambda_O^- - 1)}{1 + \lambda_O^-} \right]$
Half sib	$\frac{(2\psi - 1)}{4(p + 1)} \left[\frac{p(\lambda_O - 1)}{1 + \lambda_O} - \frac{(\lambda_O^- - 1)}{1 + \lambda_O^-} \right]$
First cousin	$\frac{\{4\psi(1 - \theta)^2 + 2\theta^2 - 1\}}{8(p + 1)} \left[\frac{p(\lambda_O - 1)}{3 + \lambda_O} - \frac{(\lambda_O^- - 1)}{3 + \lambda_O^-} \right]$

and the total expected cost could be reduced by taking this into account. On the other hand, type I error is inflated, which implies that an increase in cost might be required; but the increase in α is small in a genome-wide scan and we may therefore ignore it.

Another modification of DESPAIR is to allow for the analysis of discordant relative pairs. When only discordant sib pairs are used, the deviation from the null hypothesis mean proportion of alleles shared IBD at the first marker on either side of a disease locus δ is as given by Guo and Elston [3], and the deviation at the j -th marker on either side of the i -th disease locus is similarly $\delta_{ij}(\theta_{ij}) = \delta_j(\theta_{ij}) = (2[\theta_{ij}^2 + (1 - \theta_{ij})^2] - 1)\delta$, as given in (3), but now δ is a function of λ_O^- and λ_S^- , offspring and sib-

Table 2. Relative sample size needed to approximate the number of non-full sibling pairs^a

Lambda	Number of markers														
	300					400					500				
	1.2	1.5	2.0	2.5	3.0	1.2	1.5	2.0	2.5	3.0	1.2	1.5	2.0	2.5	3.0
Relative pair type ^b															
S	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
G	1.52	1.25	1.00	0.85	0.75	1.56	1.28	1.02	0.86	0.77	1.58	1.30	1.04	0.88	0.77
H	1.71	1.40	1.12	0.96	0.85	1.70	1.40	1.11	0.95	0.84	1.69	1.40	1.11	0.95	0.84
A	1.86	1.53	1.23	1.06	0.94	1.81	1.49	1.19	1.01	0.91	1.79	1.47	1.17	1.00	0.89
C	2.39	1.88	1.42	1.15	0.99	2.34	1.84	1.39	1.13	0.97	2.30	1.82	1.38	1.12	0.95

^a The number of relative pairs required has been divided by the number of sibpairs for the same number of markers and the same values of λ . To obtain these estimates we used the following DESPAIR parameters: mean test statistic power = 0.80, α = 0.05, R = 100, genome length = 36 M, linked distance = 0.2 M, screened proportion = 0.5 and heterogeneity proportion = 0.3. Except for where the entries are italicized, where $k = 0$ led to minimum cost, in all situations minimum cost required $k = 1$.

^b S = Sibling; G = grandparental; H = half sibling; A = avuncular; C = cousin.

ling non-recurrence risks (i.e. the probability that the relative of an affected person is unaffected, divided by the population prevalence of the disease), respectively. When contrasting affected and discordant relative pairs, which is far preferable [7], the deviation Δ from the null hypothesis of no difference in mean proportion of alleles shared IBD was expressed by Guo and Elston [3] as a function of both the recurrence and non-recurrence relative risk ratios, as well as the recombination fraction between the marker and the disease locus (see table 1).

Other parameters to be specified include the ratio (R) of the cost of recruiting (and phenotyping) one person to the cost of measuring one marker genotype, the number of disease loci assumed to exist, the genome length in Morgans, and the 'linked distance', or the maximum distance (g) between the disease locus and a marker locus to be considered linked, in Morgans.

The new version of DESPAIR also converts an average marker polymorphism information content (PIC) value into the equivalent linkage information content (LIC) value [8, 9] appropriate for the type of relative pair, so that it is no longer necessary to input f , which in the original version of DESPAIR was the parameter that allowed for the use of less than fully informative markers.

It is possible to specify the option to include the screening cost (yes or no option) to obtain the number of pairs for study and, if so, the percent of screened persons (rs) who enter the sample must be specified. Where discordant pairs are included in the study together with concordantly affected pairs, the ratio of affected/discordant rel-

ative pairs must be specified. Regarding the number of markers, the maximum number of second-stage markers to be considered per side of each significant first-stage marker should be included and the number of stage 1 markers that will be used may be optionally specified. Finally, once the desired parameter values are entered in the design page and the user selects 'submit', the DESPAIR output is displayed providing the total cost and the number of relative pairs for each value of k of second stage markers per side of a significant first stage marker, up to the specified maximum number.

Inherent in the design and conduct of an initial genome-wide linkage scan, many interesting genomic regions are not completely covered by genome-wide marker panels. In such situations where the spacing of markers does not adequately reflect the complete genetic variation at specific candidate loci of interest, one cannot be sure whether or not to rule out linkage to those candidate loci. This quandary led us to use DESPAIR to estimate the approximate power attained in a completed genome-wide linkage study to detect linkage to a candidate locus where we had used microsatellites, with the aim of determining whether typing additional markers on this same sample would be worthwhile. We will illustrate the application of this novel use for DESPAIR to our recent genome-wide linkage study of families with colorectal and breast neoplasia.

In brief, our study sample consisted of 159 full sibling pairs from 33 nuclear families; 62 concordantly affected pairs, 78 discordantly affected pairs, and 19 concordant-

600					700				
1.2	1.5	2.0	2.5	3.0	1.2	1.5	2.0	2.5	3.0
1	1	1	1	1	1	1	1	1	1
1.60	1.31	1.04	0.89	0.79	1.78	1.46	1.17	1.00	0.78
1.69	1.39	1.10	0.95	0.84	1.69	1.39	1.10	0.95	0.83
1.77	1.45	1.15	1.00	0.87	1.75	1.44	1.15	0.99	0.87
2.28	1.80	1.35	1.12	0.95	2.27	1.79	1.35	1.11	0.95

ly unaffected pairs [10]. Each pedigree included 2 cases of advanced polyps or colon cancer and at least 1 primary breast cancer. In this study, we were interested in linkage of the colon and/or breast cancer phenotype to the *CHEK2* gene on chromosome 22, a candidate gene for breast cancer [11]. We had conducted the initial genome-wide linkage analysis using the model-free methods implemented in the SIBPAL program in the S.A.G.E. program package [12], and using the Haseman-Elston method we found insufficient evidence for linkage to the microsatellite markers nearest to the *CHEK2* locus. Hence, our primary goal was first to estimate the power we had to detect linkage to this locus, given the sample and the two *CHEK2*-flanking markers, D22S689 and D22S685, and then identify whether we could materially improve power by typing second stage markers in the same sample without recruiting more individuals into the study.

To estimate the approximate power we already had to detect linkage to the *CHEK2* locus given our existing sample of sibling pairs, we used the DESPAIR software and its parameters slightly differently than originally intended. We selected the ‘full sibs’ and ‘both’ option for concordance type. We set alpha at 0.05 and power at 0.9, used the approximate method and the mean statistic, this being appropriate for a test that contrasts the allele sharing between concordant and discordant pairs [7]. We set the *Ratio of recruiting/marker assay cost* at 0 (since we were not planning on recruiting more persons), the *Number of disease loci* at 1, and the *Linked distance* at 0.1. The average PIC for the 2 flanking markers is 0.73. For *Het-*

erogeneity proportion (which we discuss later) we used 0. The *Percent of screened persons* in the sample was set at 1.0, since all subjects were genotyped. The *Ratio of affected/discordant pairs* was set at 0.8, and the *Maximum number of second-stage markers per side of significant first-stage marker (k)* was 0, because at this point we were trying to estimate power in the original scan. Finally, for the *Number of first-stage markers*, we used 2 to represent the two *CHEK2* flanking markers and a *Genome length* of 0.1 Morgans because that is the approximate distance between our two markers.

Using these values, we conducted a number of DESPAIR trials in order to determine the offspring and sibling recurrence and non-recurrence risks that gave us our 140 pair sample size available when power was set at 90% and $k = 0$. This provided two different sets of λ that would reflect the power of our study, and so we could confirm later that either set would lead to the same answer to our question of interest. We found we had 90% power to detect offspring and sibling recurrence risks of 1.23 with non-recurrence risks of 0.65, or offspring and sibling recurrence risks of 1.47 with non-recurrence risks of 0.7, assuming equal risks for siblings and offspring.

We now used these values of λ_s and λ_o and changed the maximum value of k (number of second-stage markers per side of significant first-stage marker) to 3 to examine the increase in power to be obtained by typing additional markers in the second stage. By typing two additional markers ($k = 1$) we could gain approximately 2.2% power for a cost of 2,430 units (the unit is the cost

of typing one marker on one person) or equivalently detect a smaller recurrence risk (λ_s) of approximately 1.39 – a decrease of only 0.08. We concluded that a second stage was not worthwhile here, for a gain of only 2.2% power (or to detect a λ_s 0.08 smaller than we could already detect in the original scan); therefore no additional fine mapping was advisable for this particular genomic region if this was the only sample available.

In many situations when the incidence of disease is rare, it may be more practical and more likely that relative pair types other than sibling pairs are available for study. A current limitation to utilizing other pair types for linkage studies is the lack of power estimators for model-free linkage analysis. To study mathematically how the number of required full sibling pairs needed for a linkage study equates to the numbers of other relative pair types in this two-stage analysis requires the solution of complicated equations. We therefore give empirical results that show how the number of non-sibling pairs compares to the number of full sibling pairs under a variety of scenarios. We evaluated the equivalency of grandparental, avuncular, half-sibling, and first cousin pairs to full sibling pairs by performing several DESPAIR trials to calculate the ratio of non-full sibling relative pairs to full sibling pairs needed according to recurrence risk and the number of markers used. The relative number of full sibling pair equivalents for affected grandparental, half-sibling, avuncular, and cousin pairs when $\lambda_s = \lambda_o$ are given in table 2. The results demonstrate that, whereas the number of relative pairs needed changes drastically with the specified alternative (λ), it hardly changes with the number of markers. For example, the number of half-sibling pairs required is ~ 1.71 times the number of full sibling pairs required when offspring and sibling recurrence risks (λ_o and λ_s) are both 1.2, regardless of the number of markers used, and the number of grandparental pairs, avuncular pairs, and cousin pairs needed is always about 1.52, 1.86, and 2.39 times the number of full sibling pairs, respectively. For these comparisons we only considered the case $\lambda_o = \lambda_s$, because λ_s is irrelevant for the other pair types.

The ratios of non-full sibling pairs to full sibling pairs are not influenced by use of the exact method instead of the approximate method, nor change with the linked distance. Similarly, increasing the heterogeneity proportion increases the overall sample size for all pair types, but the ratio of pairs needed does not change.

It should be noted that there are two ways to allow for locus heterogeneity if we wish to assume that in a small region of the genome one locus accounts for only a pro-

portion of the genetic component underlying a disease. One way is to allow for a proportion h of the pairs in the sample to be segregating elsewhere, i.e. h is the unlinked proportion; in this case one would assume that the linked locus is a highly penetrant locus. Alternatively, assuming all the loci, with risk ratios $\lambda_1, \lambda_2, \lambda_3, \dots$, act multiplicatively, we can use the risk ratio(s) that result when we equate $1 - h$, the linked proportion, to $\lambda/\prod_i \lambda_i$ or $\log \lambda/\sum_i \log \lambda_i$, the former being more conservative than the latter.

Finally, a problem for future research is how to incorporate non-independent pairs from the same pedigrees. If we consider all the relative pairs in a 3-generation pedigree – such as sibling, avuncular, grandparent-grandchild and first cousin – as independent, there will be serious double counting of the linkage information. Using the sibpair equivalencies we have demonstrated here and the results given by Schaid et al. [13], we anticipate that it should be possible to extend DESPAIR to design a linkage study to have the desired power when pairs are not independent, although, in view of the genotyping platforms now available, future studies (except as illustrated in this report) will have little need of two-stage designs.

Acknowledgements

This work was supported in part by a U.S. Public Health Service resource grant RR-03655 from the National Center for Research Resources, research grant GM-28356 from the National Institute of General Medical Sciences, cooperative agreement DK-57292 from the National Institute of Diabetes and Digestive and Kidney Diseases, the Cedars-Sinai Board of Governors' Chair in Medical Genetics, and NIH R25 CA94186, CA82901 and K23CA81308. Some of the results of this paper were obtained by using the program package S.A.G.E., which is supported by a U.S. Public Health Service Resource Grant (RR03655) from the National Center for Research Resources.

Web Resources

S.A.G.E. DESPAIR program, <http://darwin.cwru.edu/despair>.

References

- 1 Elston RC, Guo X, Williams LV: Two-stage global search designs for linkage analysis using pairs of affected relatives. *Genetic Epidemiology* 1996;13:535–558.
- 2 Guo X, Elston RC: Two-stage global search designs for linkage analysis i: Use of the mean statistic for affected sib pairs. *Genetic Epidemiology* 2000;18:97–110.
- 3 Guo X, Elston RC: Two-stage global search designs for linkage analysis ii: Including discordant relative pairs in the study. *Genetic Epidemiology* 2000;18:111–127.
- 4 Blackwelder WC, Elston RC: A comparison of sib-pair linkage tests for disease susceptibility loci. *Genetic Epidemiology* 1985;2:85–97.
- 5 Ziegler A, Boddeker I, Geller F, Muller HH, Guo X: On the total expected study cost in two-stage genome-wide search designs for linkage analysis using the mean test for affected sib pairs. *Genetic Epidemiology* 2001;20:397–400.
- 6 Risch N: Linkage strategies for genetically complex traits. II. The power of affected relative pairs. *American Journal of Human Genetics* 1990;46:229–241.
- 7 Elston RC, Song D, Iyengar SK: Mathematical assumptions versus biological reality: Myths in affected sib pair linkage analysis. *American Journal of Human Genetics* 2005;76:152–156.
- 8 Guo X, Elston RC: Linkage information content of polymorphic genetic markers. *Human Heredity* 1999;49:112–118.
- 9 Guo X, Olson JM, Elston RC, Niu T: The linkage information content value of polymorphism genetic markers in model-free linkage analysis. *Human Heredity* 2002;53:45–48.
- 10 Daley D, Lewis S, Platzer P, MacMillen M, Willis J, Elston RC, Markowitz SD, Wiesner GL: Identification of susceptibility genes for cancer in a genome-wide scan: Results from the colon neoplasia sibling study. *American Journal of Human Genetics* 2008;82:723–736.
- 11 Meijers-Heijboer H, Wijnen J, Vasen H, Wasielewski M, Wagner A, Hollestelle A, Elstrodt F, van den Bos R, de Snoo A, Fat GT, Brekelmans C, Jagmohan S, Franken P, Verkuijlen P, van den Ouweland A, Chapman P, Tops C, Moslein G, Burn J, Lynch H, Klijn J, Fodde R, Schutte M: The chek2 1100delc mutation identifies families with a hereditary breast and colorectal cancer phenotype. *American Journal of Human Genetics* 2003;72:1308–1314.
- 12 S.A.G.E. [2009] Statistical Analysis for Genetic Epidemiology, Release 6.0.0:<http://darwin.cwru.edu/>.
- 13 Schaid DJ, Sinnwell JP, Thibodeau SN: Testing genetic linkage with relative pairs and covariates by quasi-likelihood score statistics. *Human Heredity* 2007;64:220–233.