Linkage Newsletter

Vol. 6 No. 2 July 1992

Published by Jurg Ott, Columbia University, New York[1].

Editorial Assistant: Katherine Montague
Fax: +1-212-568-2750 Tel. 212-960-2507
e-mail: OTT@NYSPI.BITNET

Postal address: Columbia University, Unit 58
722 West 168th Street, New York, NY 10032

EDITORIAL

This summer issue of the newsletter is rather short. It is again being distributed via e-mail (to the GENETIC-LINKAGE news-group of the BIOSCI network and to individual subscribers) as well as by postal mail. Hope everyone has a nice summer!

LINKAGE COURSES

The following linkage courses are scheduled:

Zurich (Advanced course): October 19-23, 1992, in the Computer Center at the University of Zürich, Irchel campus (see full-page course announcement attached).

New York (Advanced course): January 11-15, 1993, in the microcomputer classroom of the Health Sciences library, Columbia University. Course fee $100 (supported by National Center for Human Genome Research; only a small number of participants from outside the U.S. can be admitted). Registration is open; topics covered are as in the Zürich course. A formal course announcement will be made later.

In the spring of 1993, there will also be introductory courses in New York and Zürich but dates have not yet been fixed. For information and application forms for any course, please write (preferably by fax) to Katherine Montague, course coordinator, at the address given above.

[1] Support through grant HG00008 from the National Center for Human Genome Research is gratefully acknowledged.

SOFTWARE NOTES

Bug in Vax version of MLINK?

Joseph Terwilliger recently noticed the following discrepancy between the Vax version and the PC version of the MLINK program. Consider two parents and their child who are typed for two 2-allelic markers. Both parents are heterozygous, 1/2, at each locus and the child is homozygous, 1/1, at each locus. There is complete linkage disequilibrium, P(1-1) = P(2-2) = 0.5, and P(1-2) = P(2-1) = 0. Without allelic association, of course, this family is uninformative for linkage. Under complete disequilibrium, however, given the offspring's phenotype, parental phases are known and lead to two nonrecombinations. Thus, the lod score is given by $Z(\theta) = \log[4(1-\theta)^2]$, which has a maximum of 0.60 at $\theta=0$. This is also the result furnished by the PC version of MLINK, but the Vax version reports zero lod scores at any θ. We are investigating this matter and will report results as soon as possible.

Undeclared loops

In the LINKAGE programs, when loops (marriage or consanguinity) are present in a pedigree, this must be specified in the pedigree file (the safest method is to use the MAKEPED program). If a loop remains undeclared, depending on the type of loop, the analysis programs may terminate with an error or, worse, they may appear to terminate normally but give incorrect results. Xiaoli Xie has now written a program to detect loops (Xie X, Ott J [1992] *Am J Hum Genet*, abstr, in press). It is best run after MAKEPED in which case it catches any loops the user failed to declare. This LOOP program is based on the depth-first search algorithm in graph theory and is freely available.

Simulating under heterogeneity in SLINK

Both the SIMLINK (M. Boehnke) and SLINK (D. Weeks) programs allow for simulating pedigree data under heterogeneity, that is, with a given proportion of families simulated without linkage between trait locus and marker loci. In the analysis, however, there is a major difference between the two programs: The SIMLINK program analyses the data under heterogeneity while the analysis programs of the SLINK package do not. Some users have previously been unaware of this and have analyzed the data under homogeneity even though they had been generated by SLINK under heterogeneity (the resulting expected lod scores are too small). The program ELODHET was developed to allow analysis under heterogeneity for data generated by the SLINK program. It is part of the current package of SLINK.

Change in the LINKLODS program

The LINKLODS program, which comes with the PC version of LINKAGE, calculates lod scores for individual families from the output of MLINK or LINKMAP. Dr. Chantal Mérette recently pointed out that occasionally the LINKLODS program gives incorrect "total" lod scores. This occurred when a large number of families was analyzed and the total log likelihood was smaller than the program constant *lowlod*, which previously was set to -500. This constant has now been changed to a value -10000 so that the error should no longer occur. The current program version is 1.70 (6 July 1992).

Linkage analysis with highly polymorphic markers

Large numbers of alleles can pose problems in linkage analysis. Various exact and approximate ways of overcoming these problems have been proposed. Before discussing an overview of these possibilities, I would like to ask the readers if anyone has experience with the URP program (Michael S. Braverman: "An algorithm to improve the computational efficiency of genetic linkage analysis," *Comp Biomed Res* **18**, 24-36, 1985). Please let me know ─ I'm sure many researchers will be interested in this topic.

CORRECTIONS IN *ANALYSIS OF HUMAN GENETIC LINKAGE*

Below, the currently known corrections to this book (J. Ott, 1991, Johns Hopkins University Press, Baltimore) are listed.

Page 14, line 4 up: Assumption (2) is sufficient for that statement; (2) implies (1).

Page 18, line 8: Replace (1.3) by (1.2).

Page 38, Problem 2.2: Replace *200 cM* by *100 cM*.

Page 44, line 8 below table 3.1 should read: "Generally, for phase known data, if T=k/n is the value of...". Also, line 12 should read: "Since T is unbiased, ..."

Page 47, lines 5-8: These two sentences are clearer when worded as follows: "Consider now our previous hypothetical example of one recombinant and four nonrecombinants and test H_0:θ=½ against H_1:θ=0.1. For these data, the likelihood ratio is calculated as $T_{obs} = [0.1\times(0.9)^4]/(0.5)^5$."

Page 48, line 16: Replace $A\approx(1-\beta)$ by $A\approx(1-\beta)/\alpha$.

Page 59, line 3 from the bottom: ..., $P(0\leq\theta<½) = 1/22$, ...

Page 60, lines 6 and 7 should read: The *i*th segment (i = 1..s), of length b_i, then contains the likelihood ratio, $L^*(\theta_i)$, where $b_1 = ½(\theta_2+\theta_1)$, $b_i = ½(\theta_{i+1}+\theta_i)-½(\theta_i+\theta_{i-1}) = ½(\theta_{i+1}-\theta_{i-1})$, $b_s = 0.5 - ½(\theta_s+\theta_{s-1})$; $\Sigma b_i = 0.5$.

Line 17 should read: 52.672, resulting in a value of 0.71 for Smith's (1959) posterior...

Table 4.1: The values of b_i for i=1 (now 0.025) and i=2 (now 0.050) should be 0.030 and 0.045, respectively. This way, they are consistent with the definition of the b_i's further up on page 60.

Page 34, lines 17 and 18 up are clearer when formulated as follows: "... often used before linkage analysis as a preliminary test of paternity."

Page 45, lines 12 and 13 should be phrased more exactly as follows: "..., which allows the calculation of approximate confidence intervals from asymptotic variances... ."

Page 68, last line before section 4.5: Replace 11.7 by 11.6.

Page 74, line 3: Replace Z() and $Z(_f)$ by $Z_1(_m)$ and $Z_2(_f)$.

Page 75, line 5: Replace $(1-\alpha_1)^n$ by $(1-\alpha_1)^g$.

Page 92, line 3: Replace *A1* by *A2*.

Page 93, table 5.3, line i=4: Replace AB-22 by AB-11.

Page 101, after equation (5.15): Replace $1/[n\times i(r)]$ by $1/[n\times i(r)]^{½}$.

Page 101, line 6 in section 5.9 should read: "type 1 is a recombinant under one of the parental phases (phase I, say) but a nonrecombinant under the other, ..."

Page 117, lines 21-23: The last sentence in this paragraph should read: The second child has genotype 121/222 or 122/221, *each of which requires at least one recombination in the father or the mother*.

Page 137, first line, should read: ...between the loci C and *D*.

Page 139, Table 6.10, line R: Replace "$444\theta_{BC}$" by "$444\theta_{AB}$".

Page 148, line 11: Replace f_{dd} by f_{DD}.

Page 149, table 7.1, line d1/d1: replace ½ by *½r* for P(g;r) (as on the line above it).

Page 216, Problem 9.2, line 2: Replace "table 9.6" by "table 9.7".

Page 250: The last sentence of the top paragraph contains a typo: -2 should be 2, and Z(α,x) was not defined. For better clarity, the last two sentences in that paragraph should read: "In practice, this means that one evaluates Z(,x) at each map position, *x*, where Z(α,x) is analogous to (9.9) with θ_1 replaced by *x*, and is determined by the maximum of Z(α,x) at the given *x* value. Only those points *x* are then excluded for which Z(,x)<2 and Z(x)<-2, where Z(x) is the lod score under homogeneity."

Page 268, Solution 9.2, line 2: Replace "table 9.6" by "table 9.7".

Page 270, line 1: Replace _ by _. Line 3: Replace "with that mutation" by "without that mutation".

Page 279, ref. Hall et al. (1990): Replace "Anserson" by "Anderson".

Page 294, line 2 up should read: "...tetraploid..."

Page 302, Support interval: Replace 110 by 55.

Advanced Linkage Course

October 19-23, 1992 University of Zürich, Switzerland (Irchel Campus Computer Center, 12 IBM PS/2s), Proff. Eric Kubli (Zürich) and Jürg Ott (New York).

Tuition for the 5-day course is $600.
Maximum number of participants is 12.
Application deadline: August 25, 1992.

TOPICS include: Theory of linkage programs. Practical exercises using the LINKAGE and other programs. Handling of inbreeding loops, age-dependent penetrance, and sex-specific recombination fractions. Problems of interference in multipoint mapping. Models of genetic heterogeneity. Calculation of genetic risks, also under allelic and nonallelic association. Linkage analysis with quantitative trait loci, biological covariates, and pseudoautosomal loci. Computer simulation methods. Gene mapping in CEPH reference families.

Participants must be familiar with IBM PCs or compatible microcomputers. Extensive experience with a linkage program and/or an excellent background in statistical genetics and linkage analysis are additional criteria for admission.

To obtain further information and an application form, contact:

Katherine Montague, Course coordinator, Columbia University, Unit 58, 722 West 168th Street, New York, NY 10032
Fax +1-212-568-2750 Tel. 212-960-2507