

Linkage Newsletter

No. 1

August 1987

PREFACE

For several years now, some of us working in the area of linkage analysis have been playing with the idea of circulating a newsletter as a means of rapidly exchanging the kind of information that is of importance for one's work but that is not suitable for publication in a regular journal. While previous attempts at issuing a genetics newsletter have not been very fruitful, this represents at least an initial effort at creating and maintaining a viable means of communication among researchers interested in linkage analysis. Presently, this newsletter is mailed to about 200 researchers in all continents.

Below, I propose a general outline for the newsletter; any suggestions from readers will be gladly received.

OBJECTIVES OF NEWSLETTER

1. Basically, the newsletter should readily disseminate pieces of information relevant to genetic analysis that cannot usually be published in scientific journals, where emphasis is on linkage analysis. This comprises, for example, proposals for new methods, notes on computer programs, computer hardware, meetings, experiences with analysis methods and -- last but not necessarily least -- gossip.
2. Readers are encouraged to submit articles or notes to be published in the newsletter, and I will make every effort at quickly including such contributions. I will not exercise any censorship, except that advertisements for commercial products may have to be paid for (fee yet to be determined). When announcing computer programs you will have to specify at what cost they are available.
3. A Question and Answer section is planned in which I will discuss questions I have been asked and what I consider as meaningful answers. Again, contributions from readers to this section are welcome.
4. For electronic communication among geneticists worldwide, the Bitnet/EARN network presently appears most suitable and least expensive (actually, it is free when you have an account at a computer that is a Bitnet node). At regular intervals,

therefore, Bitnet addresses of geneticists will be published in the Newsletter.

5. Funding for such an undertaking is difficult to obtain. Therefore, presumably, I will be forced to ask for a modest contribution, but this and one or two subsequent issues are distributed free of charge. Please feel free to copy the Newsletter and distribute copies to anyone interested. There is no copyright to any article or note appearing in the newsletter. If the Newsletter was not addressed to you but you are interested in receiving it directly from me, please send me your address.

SOFTWARE NOTES

Recently, Dr. Eric Lander kindly gave me a demonstration of his MAPMAKER program. It is written in C and presently runs under Unix on the HP9000 minicomputer. It is operated through a shell of commands that make it very easy to select markers, try various gene orders, maximize the likelihood under a given order, etc. I was impressed by the high degree of user-friendliness of the program. One may remember from Dr. Lander's presentation at the Philadelphia meeting last fall that MAPMAKER allows analyzing large numbers of codominant loci in families of two or three generations.

Regarding the Linkage program package (Lathrop et al.), a new version (4.5) is being distributed. A shell to allow for user-friendly operation will be added in the next few weeks. The whole package will be presented in Paris during the Human Gene Mapping workshop, in the CEPH building. Version 4.5 has faster execution speed than previous versions. The corresponding programs for small family structures allow analyzing many more loci. Again, Turbo versions will be available for the IBM PC and AT. A separate program will be available to carry out error checks, for both datafile and pedfile, and crosschecks between datafile and pedfile.

We have recently been using the SIMLINK program of Dr. M. Boehnke quite a bit. It approximates the expected lod score for a pedigree with given disease genotypes but as yet unknown marker genotypes, that is, it indicates the usefulness of a pedigree for linkage analysis prior to collecting marker data. We apply it to determine which of several pedigrees is most suitable for typing; also, the program seems a nice tool for grant proposals. Presently, it does not allow for reduced penetrance so that, for example, with X-linked inheritance, a mother with an unaffected son cannot be a possible gene carrier.

QUESTIONS AND ANSWERS

Q1: I have a marker locus with 2 codominant alleles. For some pedigree members, however, I can only determine whether allele 1 is present or not. In LIPED, no problem exists to accommodate such phenotypes, but is there such a possibility to allow for both kinds of phenotypes (codominant and dominant) in a binary factor locus in the Linkage programs?

A1: Not directly, as far as I know. However, a simple general method of converting any locus in LIPED format to Linkage format is the following. As an example, consider a marker with 2 alleles, A and B, and phenotypes as in the following table:

Geno- type	P h e n o t y p e s				
	Dominant		Codominant		
	A+	A-	A	AB	B
A A	1	0	1	0	0
A B	1	0	0	1	0
B B	0	1	0	0	1

The body of this table represents what would go into LIPED as part of the locus description. In the Linkage programs (datafile), define this locus as an affection status locus with as many liability classes as there are columns in the table above, and enter the penetrances appropriately, that is, for each genotype, you will need to furnish 5 penetrances. Each phenotype (in pedfile) is then represented by two numbers, 2 i, where i is the column number in the table above, that is, each individual is defined to be affected, except that the unknown phenotype is coded as 0 i, i being any integer from 1 to the number of liability classes. If the marker is on the X chromosome, the penetrances given for males at the AB genotype are irrelevant. While this scheme is generally applicable, in specific cases it is possible to use coding schemes requiring a much smaller number of penetrance classes.

(JO)

Q2: In LIPED, to express presence or absence of a particular allele of an X-chromosomal marker, I use codes + and - for males, and ++, +- and -- for females. How do I do this in the Linkage programs?

A2: There is no need to distinguish these phenotypes for the sexes, in any program, since these codes refer to phenotypes rather than genotypes. Thus, you may as well use +, +- and - for either sex where, of course, +- does not occur in males. In the Linkage programs, to express presence of allele number 1 at a binary factor locus, the corresponding codes would be 1 0, 1 1 and 0 1, respectively, where, again, 1 1 would not occur in males.

PRACTICAL HINTS

This section might prove useful to many readers. Please submit examples of your own that you feel other researchers could benefit from.

H1: Running the Ilink program with new data: With new data and several marker loci, it is often useful to first find or confirm estimates of recombination fractions between adjacent marker loci, that is, to run the Ilink program for the marker loci only. However, before doing that, it is a good idea to do one run with the Mlink program to verify that the likelihood is nonzero in all pedigrees. If the likelihood is zero in one or more pedigrees, for example, due to genotype inconsistencies, then the Ilink program will still try to maximize the likelihood and will, of course, fail but only after running for a possibly very long time.

PERSONAL NOTES

Dr. Lodewyk Sandkuyl, M.D., from Rotterdam, is spending one year as a visiting scientist at Columbia University Department of Psychiatry.

Jurg Ott
Columbia University, Box 58
722 West 168 Street
New York, NY 10032

Tel. (212) 960-2504
Bitnet: OTTSR@CUVMB