Linkage Newsletter

Vol. 7   No. 2   June 1993

Editorial Assistant: Katherine Montague
    Fax: +1-212-568-2750
    Tel. 212-960-2507 or 718-960-2507 (there have been some problems recently with our
        telephone and we are not quite sure which one works best right now)
    e-mail: ott@nyspi.bitnet or jurg.ott@columbia.edu

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

EDITORIAL

Finally - we have made a major step into the electronic age. Up until now, our programs have been distributed only on diskettes. We have now set up an anonymous ftp site from which interested researchers can obtain programs and other materials (see below for details).

LINKAGE COURSES

There are still a few places available for the next Advanced Linkage Course. It will be held in Zürich (Switzerland), at the Computing Center of the University of Zürich Irchel campus, September 27 - October 1, 1993. Maximum number of participants is 12. For information and application forms please write to the address above, preferably by fax.

The Columbia University Advanced Course for the academic year 1993/4 will be held in January of 1994 but a date has not been fixed yet.

MEETING ANNOUNCEMENT

(contributed by Dr. Robert Elston)

Second Annual meeting of the INTERNATIONAL GENETIC EPIDEMIOLOGY SOCIETY, October 10-11, 1993 - New Orleans, Louisiana. The meeting will be held in the Monteleone Hotel (French Quarter) and will begin on Sunday, Oct. 10th with a joint symposium with the American Society of Human Genetics on "Genetic & Environmental Aspects of Disease." For further information and abstract forms please contact Louisiana State University Medical Center Foundation, Division of Professional Education, 433 Bolivar Street, New Orleans, LA 70112 - Phone #(504) 568-3712/FAX #(504) 568-3460.

NEW ANONYMOUS ftp SITE

An anonymous ftp site (Login: anonymous; no password) is being made available for use by people interested in genetic linkage analysis. We are in the process of uploading all our programs to this ftp site. All programs distributed by us will eventually be made available on this site. The DOS version of LINKAGE version 5.20 is now also available from us (see below). For those people who are unable to access our ftp site, we will still distribute programs on floppy disks.

The ftp site presently works fine as far as we can tell. However, because of local network congestion in the afternoon, please access the site during off-peak hours as transfer may be impossible during periods of heavy network activity. To access our site, use the command

ftp york.ccc.columbia.edu   or   ftp 128.59.97.32

When asked for a login name, enter ANONYMOUS. A password is not required. For reports on problems please send e-mail to

joe@york.ccc.columbia.edu (Joe Terwilliger).

The host computer, YORK.CCC.COLUMBIA.EDU, is running Desktop-VMS Version 1.2, so the directory syntax must be specified in VMS format, which is different from UNIX-based ftp sites. When you reach our computer, you will be in directory [anonymous], which is the root directory for the ftp site. All higher directories are off-limits, and any attempt to access such directories will block any further activity for you. That root directory contains a file, README.TXT, which contains a description of the directory structure and a list of all files available from the ftp site.

2

Each of the subdirectories in the ftp site contains a file with the directory name and extension .TXT. This file briefly describes the contents of the subdirectory. As an example, the directory [anonymous.pub.newsletter] contains the current and past issues of the *Linkage Newsletter*. Eventually, it will contain all newsletters published by us since the first issue of August 1987. Please note: We are still sending out the newsletter by e-mail. If you want to receive it by e-mail as soon as a new issue is ready, send a message to:
jurg.ott@columbia.edu

When you logon, you will notice that there are some files in the root directory. These are all account specific files, which are read/write protected. You must go to the subdirectory *PUB.DIR*, with VMS syntax *[anonymous.pub]*. You can get there by typing *CD [.pub]*, for example, if your computer is running UNIX ftp software. A directory name preceded by a "." means a subdirectory. If the "." is not given, the computer will think you are trying to access a higher-level directory, and will block further access to the system. If, for example, you are in the directory *[anonymous.pub.linkage.dos]*, and you want to move to the directory *[anonymous.pub.linkage]*, you could enter *CD [-]*, where the "-" sign means to go back one directory level. To move to the directory *[anonymous.pub.tlinkage.dos]*, you could enter *CD [--.tlinkage.dos]*, meaning back up two levels, and then move forward to subdirectory *[.tlinkage.dos]*. If you have any further questions about VMS directory syntax, or are having difficulty navigating around the ftp site, please send e-mail to Joseph Terwilliger (JOE@YORK.CCC.COLUMBIA.EDU).

To get any files other than ASCII files (*.EXE or *.W51 files, for example, are non-ASCII), please make sure you do this in binary mode. Typically, one should type *SET TYPE BINARY* at the ftp prompt, though the exact syntax may vary from system to system. If you further must use communications software like KERMIT to get the binary files from a UNIX or VMS machine to your PC, be sure to do those transfers in BINARY format as well.

SOFTWARE NEWS

PC version 5.2 of LINKAGE

The PC version of LINKAGE version 5.20 is now also available from us. It is essentially the same as the one you may obtain from Paris with the following exceptions: We have added a few more error checks, the program manual is included not for Postscript printers but as ASCII and WordPerfect files, and the COMPILE batch file of version 5.10 has been adapted to version 5.20. We are also in the process of adapting the Turbo Pascal version to Borland Pascal such that, for example, it can run under Windows and use more memory than under DOS. This adaptation, however, may take some time to become available.

The OS/2 version of LINKAGE version 5.2 is in preparation and should be available in August. It is being adapted to NDP Pascal for OS/2, which is just now becoming available from Microway, Inc. (tel. (508 746 7341, fax (508) 746 4678). The previous, pre-release version already had excellent qualitites (see Linkage Newsletter of February 1993).

Modified LINKAGE programs in C

Robert Cottingham at Baylor College of Medicine has translated some of the LINKAGE version 5.1 programs into C and rewritten several procedures resulting in a program version with much improved execution speed. A paper describing these modifications will appear in the *Amer J Hum Genet*. The programs may be obtained from the anonymous ftp site, gc.bcm.tmc.edu. Robert Cottingham may be reached at bwc@bcm.tmc.edu.

Estimating allele and haplotype frequencies - BEWARE!

As more and more people estimate allele and haplotype frequencies using the LINKAGE programs, in particular, the ILINK program, it is important to be aware of the following two steps required for reliable estimation:

1) The program constant, *fitmodel*, must be set to *true* and the program recompiled.

2) Whenever all individuals in a pedigree are untyped for a marker, the UNKNOWN program makes those individuals homozygous for the *1* allele at that marker. While this speeds up calculations considerably, it leads to wrong results in the estimation of allele and haplotype frequencies. A possible solution is simply not to use the UNKNOWN program. We have slightly amended the UNKNOWN program such that it contains a program constant, *makehomozygous*, which should be set to *false* if no *unknown* phenotypes should be turned into homozygotes *1 1*. A compiled version of UNKNOWN with makehomozygous set to false is furnished in executable form as UNKNOWN1.EXE.

FORTRAN programs

While most of our programs are written in Pascal, a few programs are in FORTRAN. This includes PC-LIPED and NOCOM (estimation of mixture of distributions). All our FORTRAN programs are now compiled with Microsoft FORTRAN 5.1 and bound in such a way that the same executable code runs under DOS or OS/2 (the LIPED version allowing for 12 alleles runs only under OS/2).

URP program for downcoding alleles

URP is a computer program which reduces the number of alleles at a marker locus by re-labeling ("down-coding") them. The original version of this program was designed and written by Michael Braverman in PL/I computer language in the early 1980's and was

implemented on an IBM Series/1 minicomputer in Dr. Kenneth K. Kidd's lab at Yale University. In 1984 Randy Holmes converted this program into primary PASCAL language, and in 1985 Matthew Hawley adapted it to a VMS/VAX machine. We recently obtained a copy of URP from Dr. Andrew Pakstis at Yale University, and Xiaoli Xie converted it to Turbo Pascal. Xiaoli Xie tried it out on nine CEPH families with chromosome 13 marker data; her experiences with the program are given below.

The input file to the URP program has a format similar to the one used in the LIPED program. However, genotypes can only be coded with alphabet letters, so there are up to 52 different allele codes as upper and lower case letters distinguishable. Besides, the URP program only recodes one marker at a time. Since genotypes in input files to the LINKAGE programs are coded with numbers, one must replace these numbers by letters. As the URP program only recodes one marker at a time, repeatedly modifying the input file for URP is inevitable. If one wants to use the output file resulting from URP as input to a linkage program other than LIPED, one must change the letters back to numbers. Since the whole procedure is very time consuming, we only tested the program with two-point analysis.

We used nine families in the pedigree file, each family with approximately 14 individuals. There were two markers with 12 and 9 alleles, respectively. All but seven individuals were untyped. We ran MLINK (PC version) and LIPED based on the original dataset, and then reduced the number of alleles by running the URP program, which reduced the number of alleles in the first marker from 12 to 4, and in the second one from 9 to 3. We assigned equal allele frequencies to the recoded makers and then ran MLINK and LIPED again. The results turned out to be identical. The elapsed time for the recoded dataset was much shorter than for the original data set, for example, 3 seconds versus 125 seconds. Note, however, that not knowing the allele frequencies can create problems for the analysis (J. Ott, Am J Hum Genet 51:283-290, 1992).

DISCLAIMER: The above investigation was carried out because we keep receiving requests regarding the availability of such programs. We do not endorse the URP program in any way. Its application may be problematic in several respects that we are not pursuing any further. For example, with untyped individuals present in a pedigree, exact solutions become extremely cumbersome (Ott [1979] *Cytogenet Cell Genet* 25:196). Also, the algorithm applied in the URP program focuses on twopoint analysis, but cases are known in which pairwise comparisons are uninformative yet the multipoint analysis is not. The most urgent need for downcoding alleles is in multipoint analysis, but it is not clear whether the procedure employed in URP consistently works fine for multipoint analysis. For the time being, we choose to carry out downcoding by hand even though it may be tedious.

Advanced Linkage Course

Sept. 27-Oct.1, 1993 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: July 26, 1993.

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 a good background in statistical genetics and linkage analysis are additional criteria for admission.

(see reverse side for application form)

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

Application for Advanced Linkage Course by Dr. Jurg Ott   - Zürich -

Please fill out this page and submit it by FAX (our preference) or by mail.

Your name: _____

Affiliation: _____

Address: _____

_____

_____

Tel. number: _____

FAX number: _____ E-mail:_____

On a separate sheet, please describe in detail what kind of linkage analyses you have done, including which programs you used, how you used them, how many families you have analyzed, etc.

Below, please describe your experience with microcomputers and DOS (Which programs used? Know how to copy files, make directories, etc?)

Signature: _____ Date: _____

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