Technical Tips Online, Vol. 1, 1997

Linkage Designer and Linkage Reporter software for automated gene localization studies

Guy Van Camp, Wendy Balemans and Patrick J. Willems

Department of Medical Genetics, University of Antwerp, Universiteitsplein 1, 2610 Antwerp, Belgium

Keywords: Polymerase chain reaction, Cloning & sequencing

The mapping of genes responsible for hereditary diseases is routinely done by linkage analysis of large families using PCR analysis of simple tandem repeat (STR) polymorphisms covering the complete human genome. Conventionally, the typing of STRs is performed radioactively, and alleles are scored manually on autoradiographs. Recently however, the technology for analyzing STRs using different fluorescent dyes on an ABI automatic sequencer (Perkin-Elmer) has been developed. Specialized software packages called Genescan (Perkin-Elmer) and Genotyper (Perkin-Elmer) are available to perform data processing before lod scores can be calculated using the Linkage computer program (Ref. 1). This data analysis includes fragment length determination, discrimination between allele bands and shadow bands (allele calling), and grouping of allele lengths into discrete classes on the basis of the number of tandem repeats (allele binning).

We recently localized a gene for hereditary hearing loss using this approach (Ref. 2). In the process of analyzing the data, we experienced a number of problems using the most recent version (1.1) of the ABI Genotyper software. The first problem is the inefficient allele binning. As alleles with the same number of tandem repeats will often differ slightly in length, size intervals have to be defined for the different alleles before allele lengths can be replaced by allele numbers. With the ABI Genotyper software, the allele sizes of only a few individuals can be compared at the same time on the computer screen, whereas an overview of the allele sizes of all family members is needed for efficient allele binning. Allele binning using the ABI Genotyper software is therefore a laborious process which requires manual searching for the largest and smallest length for each allele by scrolling through the alleles from all individuals. The second problem is the inability of the ABI Genotyper software to produce the linkage file format, necessitating extensive manual editing of the ABI Genotyper output to produce linkage files. A third problem involves the checking of mendelian inheritance of the different alleles. ABI Genotyper can check only one nuclear family (parents and children) for one marker at the same time. Checking an extended family used in a genome search requires manual splitting of this family into many nuclear families, a tedious process which has to be repeated for every marker.

To solve these problems, we developed a new strategy and software for data analysis. A flow diagram illustrating the different procedures and new software is given (Fig. 1). We incorporated the pedigree drawing computer program Cyrillic (Cherwell Scientific), which is used in many labs. We also developed a software module (Linkage Designer) that forms an interface between the different existing computer programs. This module replaces inefficient procedures in existing software, and offers new possibilities for automated data analysis. On the basis of a file exported from ABI Genotyper, Linkage Designer allows binning of alleles, converting fragment sizes into allele numbers, mendelian inheritance checking and generation of the linkage file format in an efficient and automated way. Linkage Designer is able to import pedigree data from Cyrillic, and can export marker data to Cyrillic.

To monitor progress of genome-wide linkage mapping projects, we developed a second software module. The Linkage Reporter module requires input of two-point lod scores and genetic maps for the markers that were used. These data are combined, chromosome by chromosome, into graphical representations showing the regions where linkage is excluded. The use of Linkage Reporter is not restricted to fluorescent genome searches on automatic DNA sequencers; the

Corresponding author: gvcamp@uia.ua.ac.be

 PII: S1366-2120(08)70003-5

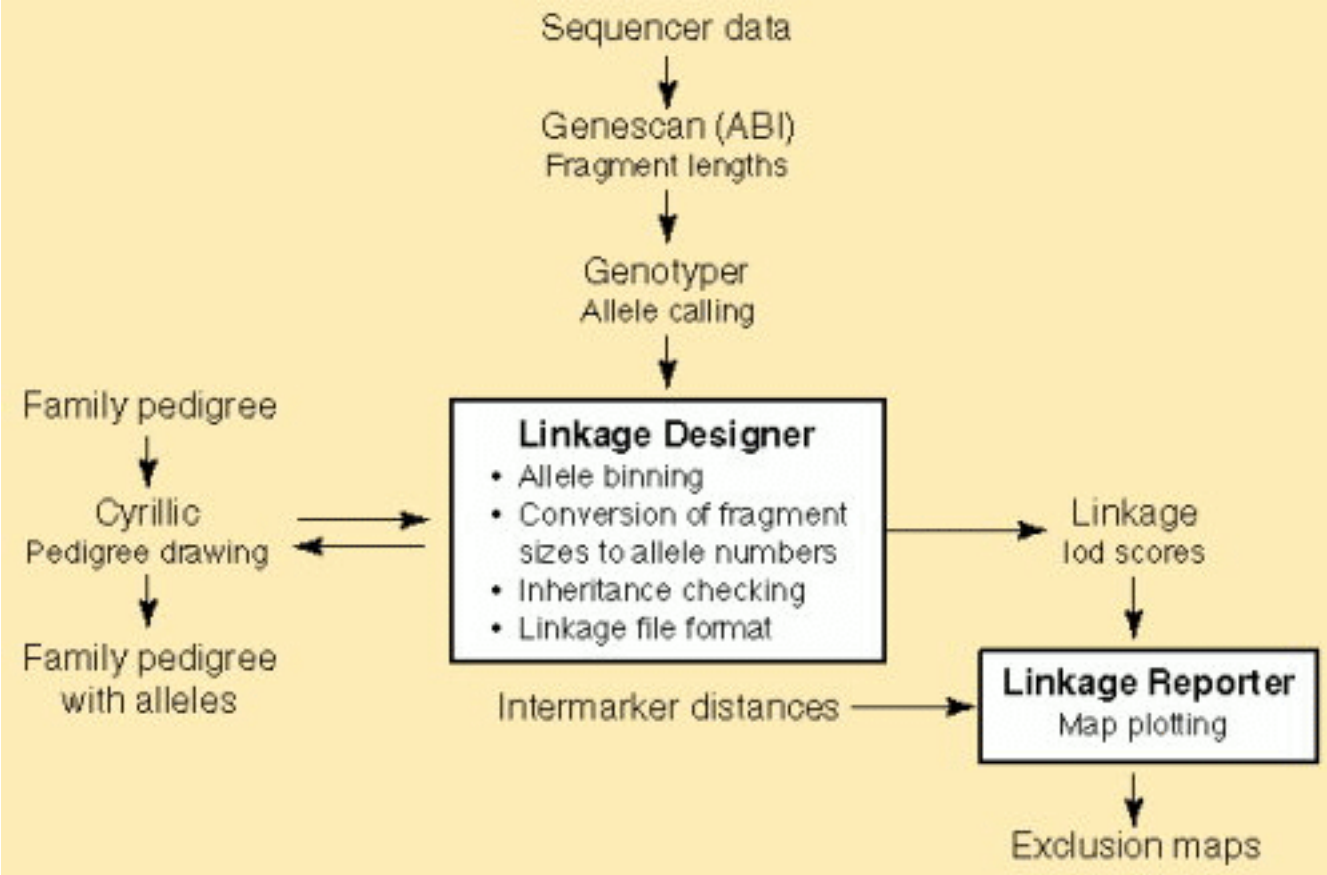

FIGURE 1. Flow diagram illustrating our strategy for data analysis.

module can also be used in traditional radioactive marker analysis.

In conclusion, our new software modules, Linkage Designer and Linkage Reporter, make data analysis in automated gene localization studies more efficient and faster. Linkage Designer and Linkage Reporter were developed in Excel for Windows v5.0 (Microsoft). The programs were tested successfully with Excel v5.0 on personal computers with Microsoft Windows version 3.1 and Windows 95, and on a Power Macintosh. Both modules contain a number of specialized worksheets, graphs, menus and associated Visual Basic code modules. Experience with Excel or comparable spreadsheet programs is recommended. The menu-operated nature of the modules should make them easy to use for anyone experienced in genetic linkage analysis. Linkage Designer and Linkage Reporter are freely available for non-profit research organizations. Both modules, including detailed instructions, can be downloaded through a World Wide Web site (http://dnalab-www.uia.ac.be/dnalab/ld.html).

Acknowledgements

This work is supported by the Belgian 'Nationaal Fonds voor Wetenschappelijk Onderzoek' (NFWO).

References

1 Ott, J. (1974) *Am. J. Hum. Genet.* 26, 588–597.
2 Camp, G. *et al.* (1995) *Hum. Mol. Genet.* 11, 2159–2163.

Products Used

automatic sequencer (Li-Cor): automatic sequencer (Li-Cor) from MWG Biotech
automatic sequencer: automatic sequencer from PE Applied Biosystems
Genescan: Genescan from PE Applied Biosystems
Genotyper: Genotyper from PE Applied Biosystems