Computer-Aided Fragment Size Determination of Single Locus DNA Probes

M. Muche

Institut für Blutgruppenserologie und Genetik, Holsteinischer Kamp 67, D-2000 Hamburg 76, Germany

INTRODUCTION

In recent years, several computer systems have been developed for fragment size determination of DNA blots. As manual evaluation is rather time-consuming, cumbersome and involves the risk of errors in measuring, calculation or transfer of data, it is obvious that a computer system can do much of the routine work.

Fragment sizes of single locus DNA probes are obtained using molecular weight markers with a known ladder of fragments. These fragments of known size are compared to the observed bands. As the migration distance of the standard size marker bands is not linear to its size, a function has to be generated, resulting in a relation between migration distance [cm] and fragment size [kb].

REQUIREMENTS FOR A COMPUTER-AIDED SYSTEM

When DNA typing was first carried out in our laboratory, the blots were evaluated using a manual direct method. With an increasing work load it became difficult to obtain fragment size values from individually drawn graphs. The first approach to that problem we took late last year was to replace the optical determination of fragment length by computerized calculation. All manually measured data was entered into a computer program that generated a function that could be used to calculate fragment sizes. It soon became obvious that some kind of a device should be introduced to make it easier to obtain centimetre values from the blot but even this rather fragmental approach proved to save a remarkable amount of time.

As an input device we decided for a digitizer. I would like to outline why this has turned out to be a way to succesfully shorten the procedure and to enhance the quality of evaluation:

The following questions can be applied on any system that is to be used for DNA fragment size determination:

(1) To what degree is manual influence required?
(2) How reliable is the system? (Keyword: repeatable results)
(3) Is a closed circuit of data provided?
(4) Is the system easy to use even if the user has little or no computer knowledge?

Advances in Forensic Haemogenetics 4
Edited by Ch. Rittner and P. M. Schneider

(5) How efficient is a system?
(6) What hardware at what cost is required?

Manual influence should be limited to what is unavoidable: The aim of any sample testing is to obtain results that are not influenced by the measuring itself. This seems to be rather difficult and it will be impossible to install a system that analyzes without any human decisions, but this influence should be restricted to clearly defined situations: We thought it tolerable that bands are marked with a mouse-like stylus on a pad, as certain decisions are to be made that a computer can hardly perform: E. g. which position of a band should be used for calculation, whether some dark dot is a band or just a bubble.

Using an optical recognition system, such as a video camera and corresponding software, has the same limitiations in this respect: There will always be cases where manual corrections or even complete manual marking is required (as for weak or broad bands). Solely manual evaluation suffers the lack of guidance: There is no plausibility check and there always exists the risk of misreading the ruler.

Repeatability is an important factor. Any two fragment size determinations of a certain sample should always result in the same value. Although a low deviation is tolerable (and will not be avoidable) this also is an indicator for the quality of evaluation. Whether or not one can obtain identical results is a question of trying: We have had a closer look at the measured migration distances of the last 40 evaluations and found differences between first and second value of between 0 and 0.08 cm, resulting in fragment size differences of up to 8 bp in the range of up to 6 kb with exceptional high deviation in the range of above 9 kb of up to 50 bp. Compared with the exactness that is achievable with manual methods this seems to be completely satisfactory. Very similar results were obtained in another laboratory that uses this system.

Every computer installation should offer what is here described as a *closed circuit*: Any data that is entered into the computer should be used for every application it is needed for: In this case the obtained data can be used for printing of a work sheet for intra-laboratory use, a blot list containing results, lists of all DNA samples examined and the calculation of probabilities in cases of disputed paternity. It should strictly be avoided that any bit of information has to be retyped or transferred manually to another program, as this allows errors to occur and drastically reduces the efficiency of the system.

Emphasis is also put on the aspect that the single steps that have to be done in order to perform a blot analysis follow a natural order. This is essential, so that the system can be *used easily* also by those who do not have any experience in dealing with computer programs or are not acquainted with the system.

The *efficiency* of the system largely depends on the following three aspects: Time needed for setting-up of hardware devices (e. g. adjustment of contrast for video systems or definiton of X and Y zero positions), duration of evaluation itself and the

degree to which advanced functions are incorporated in the system (such as population statistics, calculation of probabilities or ready-to-use printouts of phenotype or case lists).

The last aspect which is not included for scientific reasons may nevertheless prove to be important for a number of laboratories that plan to introduce a computer-aided evaluation method: *Cost*. As the digitizer seems to offer a reasonable ratio between power and price (compared to some optical recognition system), it is recommenendable even for laboratories that have a limited amount of examinations but want to maintain a high standard of accuracy.

These six aims were used as a kind of guide line in developing and installing a system that is now in use in three laboratories throughout Germany, of which a brief description is given:

PROGRAM CHARACTERISTICS

The program presented here may be used for any single locus DNA system. The probes that are to be taken in calculation can be chosen by the user. However, a set of six Collaborative Research probes is included in the package. In addition, ten different size markers can be used also allowing different size markers on one blot. The definition of a size marker includes all known fragment sizes making it easy just to click the band.

Calculation of fragment sizes is carried out using the next to marker lanes for which a result can be obtained. This result is saved for future use. Up to six results per probe may be stored for each sample. Each size determination can be confirmed or corrected by a second independent measuring. It is recommended that this second measuring is carried out by another person. A protocol listing differences is printed on demand.

Once results have been obtained these can be used for creation of a population data base or for the calculation of probabilites. A wide range of different hardcopies of data is also available.