RESTRICTION FRAGMENT LENGTH POLYMORPHISM : IMAGE ANALYSIS AND MOLECULAR WEIGHT CALCULATION WITH A SCANNER-BASED COMPUTER SYSTEM

LUCKENBACH C., LUCKENBACH A., GRATHWOHL A., RITTER H.

Insitut für Anthropologie und Humangenetik, Wilhelmstr. 27, D-72074 Tübingen
University of Tübingen, FRG

SUMMARY

Image analysis and calculation programs are essential to enable precise DNA profiling and to reduce the subjective aspects of band measurements. We present a new system, called *RFLP-MAC*, Restriction Fragment Length Polymorphism - Molecular weight Analysis and Calculation. The system consists of a scanner unit and an Apple Macintosh computer supported by a combination of a shareware and self-developed software package. Molecular weights are calculated profiling the optical density distribution of each band. The migration distances of known marker fragments can be determined by the use of different mathematical models and gel distortions are corrected by a polynomial regression algorithm. All data are stored in a relational database system. Descriptive statistics can be performed as well as allele frequencies.
Serial intergel measurements are investigated in order to check the precision of the RFLP-MAC system; these data are subjected to a standard analysis of variance (maximal 3 SD =0.030 =1.723%). Allele size distribution of 380 unrelated individuals at the locus D10S28 (TBQ7, Hae III) are presented.

1. INTRODUCTION

Highly polymorphic variable number of tandem repeat (VNTR) loci are proving increasingly profitable for genetic analysis in man, i. e. in population genetics, linkage studies, forensic medicine and paternity testing, when used in restriction fragment length polymorphism (RFLP) analysis. One problem of major concern is the accuracy and precision of fragment length calculations being obligate for further biostatistical procedures and for comparability of results.
We present a new system, specifically developed to meet these demands, called RFLP-MAC, Restriction Fragment Length Polymorphism - Molecular weight Analysis and Calculation.

2. EXPERIMENTAL TECHNIQUE

DNA-Isolation: 5 ug human genomic DNA and cell-line K562 DNA according to the protocol of Miller et al. (1).
RE-Digest: 50 U Hae III (Boehringer Mannheim / FRG).
Molecular weight Marker: NICE™ DNA Analysis Ladder (GIBCO BRL, Eggenstein/FRG) with fragment sizes ranging from 22.621 to 0.526 kb.
Electrophoresis: 0.8% agarose gel (TBE buffer), 40V, 30hrs.
Transfer: Vacuum blotting with 50 mbar
Hybridization: Alkaline phosphatase-conjugated single-locus probe TBQ7 (Promega,Heidelberg/ FRG) and the molecular weight marker probe MW100 (Zeneca Bioproducts, Frankfurt / FRG)
Detection: Chemiluminescent with Lumi-Phos 530 (Boehringer) or CSPD (Promega)

3. RFLP-MAC

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg

3.1. IMAGE-ACQUISITION

Hardware: RFLP patterns are scanned via a high resolution, 8-bit gray-scale, 300 dpi; *OneScanner*, (Apple Computer Inc.) to a *Macintosh IIci* computer with 20 MB RAM and 80 MB harddisk. Complete digital images are stored as an image file (50 KB size, PICT-Format) on the internal harddisk or external 88 MB removable cartridge.
Software: Special automatic scanning software enables to pick up even subtle changes in film or blot density and allows the user to optimize the background area.

3.2. IMAGE-ANALYSIS

Software: *Image* 1.45 (National Health Institute, Bethesda/USA). Efficient image-processing options: morphology, minimum height and area of the bands compared to the image background level, noise reduction, two-dimensional integration of the optical density within each band; this densitometric center defines each single band and is therefore the reference point for exact determination of x- and y-location measured in 0.1 millimeter-units

3.3. MOLECULAR WEIGHT CALCULATION

Software: Self-developed, running within the multi-tasking *4th Dimension* 3.0 relational database multi-user environment.
Mathematics: a) Correction of gel distortions (intragel variation): 3rd order polynomial regression algorithm, fits a curve through each corresponding triplet of the molecu-lar weight marker fragments. b) Correlation of migration and molecular weight: selection between linear, logarithmic, exponential and spline models over the full range of the molecular weight standard (global method) and between an additional interpolation model over a limited range of the standard (local method), (2-4); we use the local version of the reciprocal method.
Database management: All data are stored in a relational database system.

3.4. STATISTICAL EVALUATION

Descriptive statistics: Selection of common statistical parameters: arithmetic mean (AM), standard deviation (SD), coefficient of variation (CV), variance and confidence interval. To check these paramters we have made serial intergel measurements of the two K562 DNA fragments separated on 20 gels and hybridized to TBQ7 (HaeIII).
Population- and formal genetics: Allele frequencies of the investigated probe systems can be analysed by three methods: classical rounding, "fixed bin" (5) and "sliding window" (6) fit. All data are available in tabulated and graphical form. We investigated 380 unrelated individuals from SW Germany.

4. RESULTS AND DISCUSSION

4.1. GEL IMAGE

Fig. 1 shows the fragment patterns of Hae III restricted DNA from 16 unrelated individuals detected by TBQ7. Correction of gel ditortions by 3rd order polynomial regression algorithm requires the use of 3 tracks with molecular weight markers, running in lanes 1, 10, 19.

4.2. RFLP-MAC

The image-aquisition, image-analysis, and molecular weight calculation method described above insure consistent results from gel to gel. Determining the optical density distribution of each band results in a precise calculation of its center compared with other methods localizing bands with a ruler, micrometer or digitizing tablet. For the calculation of the molecular weight we prefer the local version of the reciprocal method, because this algorithm describes the

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg

relationship between migration and molecular weight very precise gained by using only marker fragments close to the unknown bands. The only condition for calculating with this algorithm is the use of a molecular weight marker covering the whole blot with 20-30 well-defined and distinctly separated fragments.

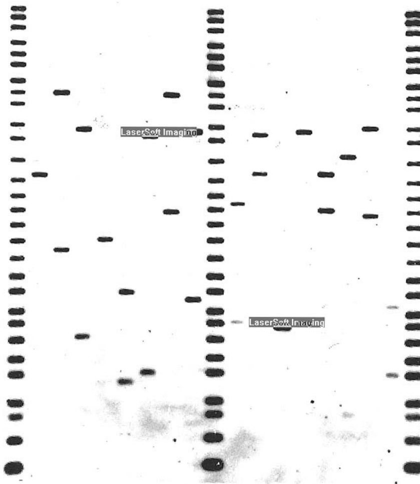

Fig.1: Hae III restricted DNA from 16 individuals, detected by TBQ7, molecular weight marker MW100 (1,10,19)

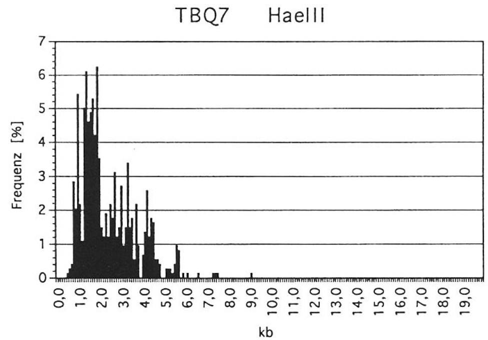

Fig. 2: Allele frequency distribution at the locus D10S28 (TBQ7/Hae III), n = 380 unrelated individuals

4.3. STATISTICAL EVALUATION
Tab. 1 summarizes the statistical results. It is apparently that low SD values correspond to precise analysis methods.

TBQ7	intergel V.	
	f1	*f2*
AM (kb)	1,742	1,173
SD	0,010	0,006
CV (%)	0,574	0,480

Table 1: Analysis of intergel variance (n=20) for molecular weight calculations of two K562 DNA fragments, detected by TBQ7 (HaeIII), (AM=arithmetic mean, SD= standard deviation, CV=coefficient of variation (%), f1,f2 =fragment 1,2)

Fig. 2 presents the allele size and frequency distribution at the locus D10S28 generated by Hae III restriction enzyme digest of DNA from 380 unrelated individuals detected by the TBQ7 DNA probe.
Summarizing the data revealed in this study using the RFLP-MAC system answer all requirements claimed by an accurate and reproducible analysis method.

REFERENCES

(1) Miller, S.A., Dykes, D.D. and Polesky H.F., Nucleic Acids Res. 1988, 16, 1215.
(2) Southern, E.M., J. Mol. Biol. 1975, 98, 503-517.
(3) Elder, J.K., Amos, A., Southern, E.M. et al., Anal. Biochem. 1983, 128, 223-226.
(4) Elder, J.K., Southern, E.M., in: Bishop, M.J. and Rawlings, C. J. (Eds.), Nucleic Acid and Protein Sequence Analysis, IRL Press, Oxford 1987, pp. 165-172.
(5) Budowle, B., Giusti, A.M., Waye, J.S., Baechtel, F.S., Fourney,R.M., Adams, D.E., Presley, L.A., Deadman, et al., Am. J. Hum. Genet.1991, 48, 841-855.
(6) Gill, P., Sullivan, K. and David, J.W., Hum. Gen. 1990, 85, 75-79.

Advances in Forensic Haemogenetics 5 (c) Springer-Verlag Berlin Heidelberg