MULTIVARIATE ANALYSIS OF ANTIGENIC VARIATION AMONG GEMINIVIRUS ISOLATES ASSOCIATED WITH CASSAVA MOSAIC DISEASE

D. Fargette[1], P.F. McGrath, J.W. McNicol, M.M. Aiton and B.D. Harrison

Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, U.K.

INTRODUCTION

Extensive work on the geographical variation among geminivirus isolates associated with cassava mosaic disease in Africa and India has been conducted for several years in the Virology Division of the Scottish Crop Research Institute and results of these studies have been summarized elsewhere (Aiton and Harrison, 1988; Aiton et al., 1988; Harrison et al., 1987). Virus isolates from 10 countries in Africa and the Indian subcontinent were tested against a panel of monoclonal antibodies (MAbs) prepared against a west Kenyan isolate of African cassava mosaic virus (ACMV; Thomas et al., 1986). On the basis of the occurrence of 17 epitopes, the isolates fell into three groups. Group A isolates came from Ivory Coast, Nigeria, Angola, South Africa and western Kenya; Group B isolates came from coastal Kenya, Malagasy, Malawi and Tanzania, and Group C isolates were from India and Sri Lanka. Within each group there was some variation: group A isolates shared at least 14 epitopes with the west Kenyan isolate, group B isolates shared 4 to 9 epitopes with it and group C isolates shared only 2 or 3. Moreover, isolates in the same group tended to share the same epitopes (Harrison and Robinson, 1988).

The numerical data obtained by ELISA in the work described above involved the reactions of several dozen virus isolates with a panel of 17 MAbs. Thus although an analysis and diagrammatic representation of the whole set of data is desirable, this is difficult in practice. A diagram summarizing the data was produced by Harrison and Robinson (1988) but this took only the qualitative information into account, each reaction between a virus isolate and a MAb being scored either positive or negative. We describe here a different system for representing the data, based on multivariate analyses of the quantitative values obtained by ELISA. First, Hierarchical Cluster Analysis (HCA) was done to put the grouping of virus isolates on a mathematical basis. Principal Component Analysis (PCA) was then conducted to establish, in space with a limited number of dimensions, the positions of individual isolates, the occurrence of clusters (groups) of isolates, and the variation within and between these groups. In addition, the grouping of different MAbs, on the basis of their pattern of reactions against the isolates, was determined and

---

[1]On secondment from ORSTOM (Institut Francais de Recherche Scientifique pour le Developpement en Cooperation)

illustrated. Finally, Stepwise Discriminant Analysis (SDA) was done to establish which MAbs comprise the minimum number needed to distinguish between the groups of isolates.

MATERIAL AND METHODS

A relatively large panel of MAbs was available, but not all the isolates were tested with the same selection of MAbs during the six-year study. Therefore two versions of the analyses were performed to minimize the amount of missing data:

(a) A smaller number of MAbs (12) and a larger number of isolates (32 in group A, 17 B and 12 C) were used. These MAbs are those known as SCR 14, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26 and 29.

(b) A larger number of MAbs (14) and a smaller number of isolates (26 A, 17 B and 12 C) were used. The MAbs included those listed above, together with SCR 28 and 32.

For both data sets an ELISA reading was taken after incubation with substrate for 1 hr at room temperature and also after subsequent incubation overnight at 4°C. The mathematical procedures used for HCA, PCA and SDA are essentially as described by Gnanadesikan (1977).

RESULTS

The three groups of isolates obtained by HCA correspond to the groups proposed previously, i.e. the West African group (A), the East African group (B) and the Indian group (C). Increasing the number of MAbs (and decreasing correspondingly the number of isolates analysed) did not change the grouping appreciably. Analyses of overnight readings allowed a better distinction between the groups than similar analyses of 1-hour readings and therefore the subsequent analyses were done on the overnight readings.

For PCA, the results were used from 61 isolates tested against a panel of 12 MAbs. The first principal component was a weighted average for all the MAbs (with low weights for SCR 18 and 20) and accounted for 61.2% of the variance. The second component had large positive coefficients for SCR 17, 23 and 26 and large negative coefficients for SCR 16, 19, 21, 22, 25 and 29. This accounted for 26.6% of the variance. The third component had large negative coefficients for SCR 18 and 20 and smaller positive coefficients for the other MAbs. This accounted for 5.1% of the variance. Therefore, PCA allows us to represent in two dimensions nearly 88% of the variance of the results contained in a 61 x 12 table. Indeed by including a third axis about 93% of the information can be represented.

Plots of the first two principal components clearly separated the three groups of isolates. However, the groups differ in the amount of variation they contain, with group B being the most variable. The variation can be attributed to several factors. It can be caused by differences in reaction intensity, such as those caused by differences between isolates in virus concentration. This kind of variation would appear preferentially along axis 1 (which represents an average of the reactions given by all the MAbs). Variation can also be attributed to the failure of some MAbs to react because the isolate lacks the corresponding epitopes. This would appear preferentially on the axis on which these MAbs have their greatest weight. Finally, variation can be

linked to differences in the relative intensity of reaction of some MAbs, such as those caused by minor changes in individual epitopes. All three kinds of variation were apparent in these studies, although most of it is caused by differences in reaction intensity and by the absence/presence of some epitopes. However, the considerable variation found among isolates in group B mainly reflects the absence of different combinations of epitopes in different isolates within this group rather than differences in concentration. Group C appears to be much more compact in plots of the first two components, but some minor variation is clearly expressed along axis 3 and is caused by variation in reactions with MAbs SCR 18 and 20 which have a strong influence on the position of points along this axis. Group A is the most compact cluster in the three dimensions, despite including the most isolates and the results of experiments conducted over a period of six years. The small variation within group A is attributed mainly to differences in reaction intensity (axis 1) and to differences in relative intensity of reactions with a few MAbs, such as SCR 19 and 22 (axis 2).

Further analysis of the correlation matrices indicates that the reactions of MAbs SCR 16, 19, 21, 22, 25 and 29 were highly correlated with each other. Similarly, the reactions of MAbs SCR 17, 23 and 26 were highly correlated. The reactions of SCR 18 and of SCR 20 were not significantly correlated with those of any other MAb. The reactions of SCR 14 were correlated significantly, although not strongly, with those of every other MAb except SCR 18 and 20. Principal Component Analysis allows a clear representation of the relationships between the reaction patterns of the MAbs, on the basis of the readings obtained, in plots of the two first principal components.

The application of Stepwise Discriminant Analysis to the results given by 55 isolates in reactions with 14 MAbs allowed us to identify the best combination of MAbs necessary to distinguish the different groups of isolates. Reactions with MAbs SCR 17, 20, 21 and 23 classified correctly 53 of the isolates. Indeed, increasing the number of MAbs did not significantly improve the classification. Only two isolates were misclassified and these exceptions can be explained by technical problems.

DISCUSSION

Multivariate analyses proved to be useful to group the virus isolates (HCA), to display the variation between and within groups of isolates (PCA), to illustrate the relationships between the patterns of reactivity of different MAbs (PCA) and to select the MAbs needed to discriminate between groups of isolates (SDA). These analyses confirm and refine previous interpretations of the data (Harrison and Robinson, 1988). They provide a clear picture of the relationships which exist and are a tool powerful enough to cope with further analyses, in which the number of virus isolates or viruses is increased. These sorts of analysis could also be helpful for assessing antigenic variation in countries or areas where the availability of MAbs is restricted and where ELISA plate-reading facilities are not available. For example, SDA would help to select the MAbs needed to study the range of variation among strains or related viruses. In addition, a grading system based on visual assessment of the strengths of ELISA reactions could be worked out which would reveal much of the antigenic variation that exists.

REFERENCES

Aiton, M.M., Harrison, B.D. (1988). Virus isolates from mosaic-affected cassava in the Ivory Coast. Rep. Scott. Crop Res. Inst., 1987: 190-191.

Aiton, M.M., McGrath, P.F., Robinson, D.J., Roberts, I.M., Harrison, B.D. (1988). Variation in Indian cassava mosaic geminivirus. Rep.Scott. Crop Res. Inst., 1987: 191.

Gnanadesikan, R. (1977). Methods for statistical data analysis of multivariate observations. John Wiley & Sons, New York, p.311.

Harrison, B.D., Lennon, A.M., Massalski, P.R., Robinson, D.J., Thomas, J.E. (1987). Geographical variation in geminivirus isolates associated with cassava mosaic disease. Rep. Scott. Crop Res. Inst., 1986: 179-180.

Harrison, B.D., Robinson, D.J. (1988). Molecular variation in vector-borne plant viruses: epidemiological significance. Phil. Trans. R. Soc. Lond. B321: 447-462.

Thomas, J.E., Massalski, P.R., Harrison, B.D. (1986). Production of monoclonal antibodies to African cassava mosaic virus and differences in their reactivities with other whitefly-transmitted geminiviruses. J. Gen. Virol. 67: 2739-2748.