Minimum Convex Polygons for the Delineation of Possible Viral Taxa

By R. E. F. MATTHEWS

Department of Cell Biology, University of Auckland, Auckland, New Zealand

(Accepted 2 February 1977)

SUMMARY

Using a plot of genome dry mass against particle dry mass for 58 virus groups, the utility of minimum convex polygons for delineating possible viral taxa has been examined.

INTRODUCTION

Minimum convex polygons have been used in various fields of biology to delineate groups based on variation in two physical characters. A minimum convex polygon is a polygon of minimal area with its vertices on data points and with no internal angles greater then 180°. They have been used very effectively to delineate lower and higher groups of vertebrates based on a log-log plot of body mass and brain mass (Jerison, 1968, 1973, 1976). On a much finer biological scale they have been used for the morphological characterization of castes within a single species, e.g. polymorphism between the castes of bees, based on a plot of wing length and abdomen width (Michener, 1974).

Two of the most basic properties of a virus that can be expressed as single numbers are its size and the size of its genome. Size of the genome 'sums up' many other properties and activities of the virus. The size of the virus itself will be related to the structural complexity of the particle. Large viruses tend to have many components and complex structures whereas small viruses have few components and a simple structure. In this paper I explore the usefulness of polygons for delineating possible viral taxa, using up-to-date data on genome size and particle size and the families, genera and groups of viruses recently established by I.C.T.V. (Fenner, 1976).

METHODS

For those viruses where the data are available for the dry particle mass and the mol. wt. of the nucleic acid I have used the type species for each approved or proposed group listed by Fenner (1976). Some of the families or genera listed by Fenner were not included because data were not available for size of the nucleic acid or of the virus. For convenience here I shall use the word ‘group’ to mean an approved or proposed genus, as well as for the groups proposed by plant virologists. Three viruses not belonging to groups listed by Fenner have been added. References for the data on these are given in Fig. 8.

On the basis of reported lipid analyses I previously placed the Baculoviruses in the group without envelopes (Matthews, 1975). Recent analysis on highly purified virus samples show 10% of lipid for potato tuber moth granulosis virus (J. Longworth, personal communication). This information together with various published electron micrographs of thin sections of virus particles, shows that the Baculoviruses must have a lipoprotein bounding membrane.
RESULTS AND DISCUSSION

In order to examine the possible use of polygons for the delineation of viral taxa, the nucleic acid mol. wt. and the virus particle mol. wt. were plotted on a log-log scale for representatives of 58 virus groups (Fig. 1). The range in particle mol. wt. and nucleic acid mol. wt. both extend over somewhat more than three orders of magnitude. As would be expected on general grounds, larger viruses tend to have larger genomes. There is no marked discontinuity on the basis of size between the largest viruses and the smallest cells.

In the past viruses have often been grouped according to the host organism. Fig. 2 shows polygons for viruses infecting bacteria, plants and animals. Although most of the larger viruses infect animals, there is substantial overlap between the three polygons.
Viral taxa: use of minimum convex polygons

Fig. 2. Polygons delimiting viruses infecting plants, bacteria and animals (all phyla). Viruses replicating in plants and insects were classed as plant viruses. ☐, Rod-shaped plant viruses; ●, all other groups; ---, polygon for plant virus groups; --, polygon for bacterial virus groups; ..., polygon for animal virus groups.

The possession of DNA or RNA has also been used as a basis for the first division in classification by several authors (e.g. Lwoff, Horne & Tournier, 1962; Lwoff & Tournier, 1971; Melnick, 1975). Fig. 3 shows the polygons for these groups. There is a substantial overlap between the two polygons.

I have suggested previously (Matthews, 1975) that a primary subdivision of virus groups might usefully be made according to whether the virus particle contains a bounding lipoprotein membrane (enveloped virus) or whether such a membrane is absent (geometric viruses). The rod-shaped plant viruses have a 5 to 10 times greater particle mass for a genome of given size than an icosahedral particle. This difference arises mainly because the icosahedral particle provides an internal protected space where part of the nucleic acid at least is not in intimate association with protein. If the rod-shaped groups are excluded, there is no overlap between the polygons for enveloped and geometric groups (Fig. 4).
Polygons delineating groups with single-stranded (SS) or double-stranded (DS) nucleic acids have a small overlap that is not due to rod-shaped virus groups (Fig. 5).

Fig. 6 shows the four divisions based on presence or absence of a bounding lipoprotein envelope and strandedness of the nucleic acids. There is little overlap between the four divisions which suggests that they should have good predictive value.

The possession of DNA or RNA has been widely regarded as an important distinguishing character between viruses. However, it appears to be unsatisfactory as a criterion for a first division (Fig. 3). Nevertheless if we use DNA versus RNA as the character for a third subdivision the seven clusters shown in Fig. 7 emerge.

In several published virus classifications the double-stranded RNA viruses (*Reoviridae*) are placed together with other RNA viruses. However on the basis of structural details and reproductive mechanisms the *Reoviridae* are coming to be regarded as a distinctive group (e.g. Joklik, 1974). The plots in Fig. 7 confirm this view. The *Reoviridae* (polygon No. 5) is well separated from both the enveloped and the geometric SS RNA viruses.
Viral taxa: use of minimum convex polygons

Fig. 4. Polygons delimiting the enveloped and geometric virus groups. Rod-shaped plant virus groups (Ⅻ) are excluded from the polygon for geometric viruses. ●, All other virus groups; ——, polygon for enveloped viruses; —— ——, polygon for geometric viruses.

The virus groups belonging to the seven divisions in Fig. 7 are enumerated in Fig. 8. There is no known example of an enveloped virus with SS DNA. There is one example with DS RNA (Φ6, Vidaver, Koski & Van Etten, 1973) which I had previously overlooked (Matthews, 1975). The enveloped single-stranded RNA groups are subdivided as suggested previously (Matthews, 1975). The subdivision of the geometric single-stranded groups has been modified from that previously suggested (Matthews, 1975) in the light of the groupings suggested by Fig. 7.

Minimum convex polygons based on a log-log plot of genome mass and particle mass might also be useful when taxa approximating to established families are being considered. For example Fig. 9 shows that 17 of the 22 geometric SS RNA groups form three clusters with common properties. (1) The established family – Picornaviridae, (2) plant viruses with rod-shaped particles, and (3) icosahedral plant viruses with genomes split between two or
Fig. 5. Polygons delimiting virus groups that contain DS, or SS nucleic acids. •, Rod-shaped plant virus groups; ○, all other groups; ---, polygon for SS nucleic acids; ----, polygons for DS nucleic acids.

more particles. I conclude that the procedure described here may be usefully employed to explore possible virus family groups, and higher taxa.

It remains to be determined whether the method will be effective for examining subdivisions within groups of family or genus size. The variation in particle mass and genome mass may be too small to give useful subgroupings. However, this question can be adequately tested only when good data are available for a sufficient number of individual viruses within a given family or genus.

In the present state of knowledge it is not possible to explain why some criteria lead to overlapping polygons and others do not; or to explain certain other features, for example, the small variation in the known geometric SS DNA viruses (Fig. 7, No. 6) and geometric DS RNA viruses (Fig. 7, No. 5) compared with the geometric SS RNA viruses (Fig. 7, No. 7). I should also stress that, at present, this form of analysis does not necessarily give any indications of evolutionary trends or relationships. The clusters of viruses with similar
Fig. 6. Polygons delimiting four clusters of virus groups using enveloped v. geometric and DS v. SS as the criteria. ➕, Rod-shaped plant virus groups; ⬢, all other groups; ———, polygon for SS geometric viruses; ———, polygon for SS enveloped viruses; ———, polygon for DS geometric viruses; . . . . , polygon for DS enveloped viruses.

properties (for example, the plant viruses with rod-shaped particles, Fig. 9, cluster 2) might indicate evolutionary relationship, but they might also reflect limitations on a particular type of structure dictated by strictly physical considerations.

I wish to thank Dr A. R. Bellamy for useful discussions and Professor F. Fenner for a copy of his manuscript prior to publication.
Fig. 7. Polygons delimiting the seven clusters that arise when the clusters in Fig. 6 are further subdivided on the basis of the kind of nucleic acid (DNA or RNA). 8. Rod-shaped plant virus groups; ●, all other groups. Enveloped: 1, DS DNA; 2, DS RNA; 3, SS RNA. Geometric: 4, DS DNA; 5, DS RNA; 6, SS DNA; 7, SS RNA.
Viral taxa: use of minimum convex polygons

A. Enveloped viruses

- **Double-stranded nucleic acids**
  - (1) Double-stranded DNA
    - Poxviridae (Vesiculovirus; fowl pox)
    - Herpesviridae (Herpes simplex)
    - Borna-viridae (Borna virus polyhedrosis)
  - (2) Double-stranded RNA
    - Cystoviridae (QX6)
  - (3) RNA
    - No DNA step in replication cycle

- **Single-stranded nucleic acids**
  - Genome = mRNA
    - Togaviridae (Sindbis; Yellow fever)
    - Bunyaviridae (Bunyamwera)
  - Genome complementary to mRNA
    - Rhabdoviridae (Vesicular stomatitis; rabies)
    - Paramyxoviridae (Newcastle disease; measles)
    - Orthomullers (Influenza)

B. Geometric viruses

- **Double-stranded nucleic acids**
  - (4) DNA
    - Iridoviridae (Tipula iridescent type 1; frog 1; Tan & McAuslan, 1971; mosquito irido [Wagner et al., 1973])
    - Adenoviridae (Adeno-type 2; CELO)
    - Papovaviridae (Shope papilloma; mouse polyoma)
    - Caulimovirus (Cauliflower mosaic virus)
  - (5) RNA
    - Reoviridae (Reo type 1; wound tumour; cytoplasmic polyhedrosis of Bombyx mori)
    - Microviridae (Coli φX174)
    - Inoviridae (Coli fd)
  - (6) DNA
    - Paroviridae (Kilkham latent rat; adenov associated type 1; Denso of Galliera)
    - Bromovirus (Bromovirus)
    - Comovirus (Cowpea mosaic virus)
    - Cucumovirus (Cucumber mosaic virus)
    - Ilarvirus (Tobacco streak)
    - Nepovirus (Tobacco ringspot)
    - Pea enation mosaic virus group (Pen enation mosaic virus)
    - Tomato necrosis virus group (Tobacco necrosis A)
    - Satellite tobacco necrosis virus (Satellite tobacco necrosis [Kassanis, 1970])
    - Tombusvirus (Tomato bushy stunt virus)
    - Turnip mosaic virus family (Alligator virus)
    - Carnovirus (Carnation latent virus)
    - Closterovirus (Bret yellows)
    - Hordeivirus (Barley stripe mosaic virus)
    - Potexvirus (Potato X)
    - Pospiviroid (Potato Y)
    - Tobamovirus (Nicotiana mosaic virus)
    - Tobravirus (Tobacco rattle)

**Fig. 8.** Virus groups in the seven subdivisions shown in Fig. 7. The names of the virus or viruses used to represent each group are listed after the group name. Data for all these viruses, except three, are given in Fenner (1976). References to these three are given in the table. Groups approved by I.C.T.V. are given in italics.
Fig. 9. Possible subdivisions among the 22 groups of geometric SS RNA viruses. The polygons delimit three genera of an established family—Picornaviridae (1) and two clusters of plant virus groups with other properties in common—rod-shaped viruses (2) and icosahedral viruses with the genome split between two or more particles (3) [alfalfa mosaic virus is included here].

REFERENCES


(Received 29 November 1976)