

Computer graphics in fungal identification

BRYCE KENDRICK

Department of Biology, University of Waterloo, Waterloo, Ontario

Received April 4, 1972

KENDRICK, B. 1972. Computer graphics in fungal identification. *Can. J. Bot.* **50**: 2171–2175.

A flexible, sequential, computer graphics image-retrieval procedure is used in constructing a simple but unorthodox key to didymosporous Hyphomycetes.

KENDRICK, B. 1972. Computer graphics in fungal identification. *Can. J. Bot.* **50**: 2171–2175.

L'auteur décrit l'utilisation d'un système mécanographique séquentiel et flexible, basé sur des images, pour la construction d'une clé simple mais non orthodoxe pour l'identification des Hyphomycètes didymosporés.
[Traduit par le journal]

Taxonomic mycologists have long been resigned to the arrival on their desks of a steady stream of fungi all insisting on instant identification. The process of placing each in its proper pigeon-hole is occasionally educational, since they spring from many states and substrates, and sometimes one discovers diamonds among the dross—hitherto unknown fungi which help to complete our catalogue of the earth's creatures. Nevertheless, the ninety and nine necessitate tedious time wasting for the expert, and he may be heard loudly lamenting as routine replaces research.

Let us look at a pair of possible escapes from this predicament. The first is to place adequate taxonomic literature in the hands of the applied mycologist, thus enabling him to identify his own fungi. The second is to train relatively unskilled assistants to screen incoming organisms and to identify them with some degree of accuracy. The two approaches are obviously not mutually exclusive, and there are compelling reasons for trying both.

There is now a large body of excellent descriptive literature on the Hyphomycetes (though as Kendrick and Carmichael (in press, 1972) have shown, hundreds of extant generic names are still insufficiently characterized). Nevertheless, many of us have had the satisfaction of publishing a supposedly adequate illustration and description of a new taxon, only to be subsequently shattered when we receive a completely different organism in the mail, masquerading under our novelty's name. At moments like that, one considers effective communication among colleagues to be impossible.

In any case, no matter how good the literature is, the painstaking applied mycologist will often want his identification checked by an acknowledged expert. Publications, especially compendia like those of Barron (1968), von Arx (1970), Ellis (1971), and Kendrick and Carmichael (in press, 1972), may reduce the influx of cultures, but they can never completely eliminate it.

It would seem logical to have the incoming cultures screened by someone who is subjected to fewer pressures than is the professional. If at least the common genera could be thus weeded out, externally imposed demands on the taxonomist's time could be sharply reduced and his research productivity increased.

With this in mind I have constructed a sample key designed for use by relatively unskilled personnel interacting in the conversational mode with a computer through its graphic display unit. In computer terms, this feasibility study is a flexible, sequential, image-retrieval procedure. This may seem a case of technological overkill, since the present key could also be adapted to the printed page (though it would be less convenient to use in that format), but if we consider the possibility of including in such a key all fungi likely to be encountered, then the flexibility and agility of the computer, with its extremely rapid access to, and display of, selected stored data, become much more valuable.

The organisms keyed out in this pilot study are the didymosporous genera of Hyphomycetes which have been isolated from soil. The data were abstracted from Barron (1968).

The basic idea behind the sequence of images used is that of passing in stages from the general-

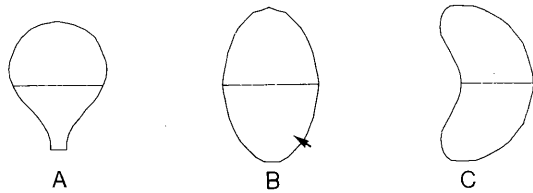


FIG. 1. Calcomp plot of the level 1 display, showing three generalized conidium shapes. The arrowhead indicates interception of image B by the light pen, leading to the display shown in Fig. 2.

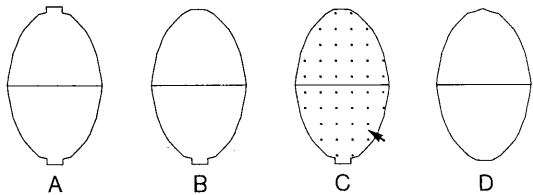


FIG. 2. Calcomp plot of a level 2 display, showing four generalized shapes, all with the basic characteristics of Fig. 1B, but differing from each other in presence or absence of apical and basal dehiscence scars and pigment. The arrowhead indicates interception of image C by the light pen, leading to the display shown in Fig. 3.

ized to the specific. The first choice the operator must make is between three different shapes (Fig. 1). The first conidium has cells of very unequal size, though it is symmetrical about the vertical axis. The second image is symmetrical about both horizontal and vertical axes, and the third is asymmetrical about the vertical axis. Pigmentation, ornamentation, appendages, and other individual quirks are deliberately ignored at this stage. The operator must choose the image which most resembles the conidium of his unknown organism, then intercept that

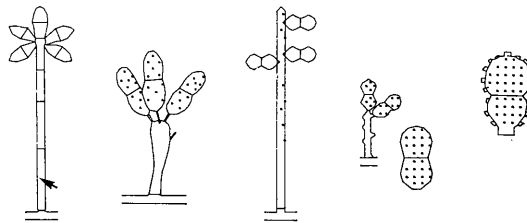
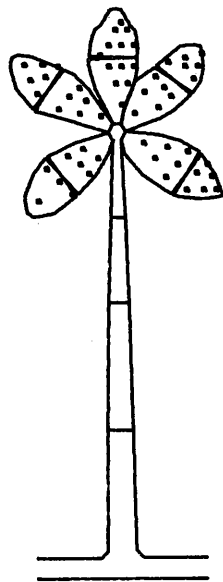


FIG. 3. Calcomp plot of a level 3 display, showing five generic images, all with the basic characteristics of Fig. 2C. The arrowhead indicates interception of the image by the light pen, leading to the display shown in Fig. 4.



**CORDANA
PAUCISEPTATA**

FIG. 4. Calcomp plot of a level 4 display, showing a more detailed image, and giving the binomial and other pertinent information as outlined in the text. This is the end product of the program.

LINNAEA 24: 129. 185
HUGHES, S. J. 1955

CAN. J. BOT. 33: 259-268

DIAGNOSTIC FEATURES -
APICAL CLUSTER OF DRY, DARK,
2-CELLED SPORES WITH VERY
DARK SEPTUM, ON SIMPLE DARK
CONIDIOPHORE. SYMPODIOCONIDIA.

DIMENSIONS -
CONIDIA 7.5-12 X 4-6.7 MICRONS
CONIDIOPHORES UP TO 170 MICRONS

OCCURRENCE -
INFREQUENT IN SOIL. COMMON ON
WOOD AND BARK OF ABIES, ACER,
BETULA, CARPINUS, PICEA.

MEDIA -
HEA, PDA

image at any point with the light pen. This automatically causes the computer to erase the first set of images and display the appropriate second set. Let us assume that the operator has chosen and intercepted image B. The set to which it leads (Fig. 2) comprises four images, basically similar in outline since they all share the basic features of image 1B, but each differing from the others in one or more features. Here, the variables are the presence or absence of apical and basal dehiscence scars, and the presence or absence of pigment (represented by dots).

If we assume that the operator selects the third (pigmented) image and intercepts one of its lines with the light pen, the computer will immediately display the next appropriate image set (Fig. 3), which consists of five much more specific images. Now, not only the conidia, but also the conidiophore and the way in which conidia are arranged on it or attached to it, are shown. It appears, then, that the third image of the second set has characteristics shared by five genera. If the operator has seen a reasonably good slide preparation of the organism (and he should perhaps have by his side a sketch with

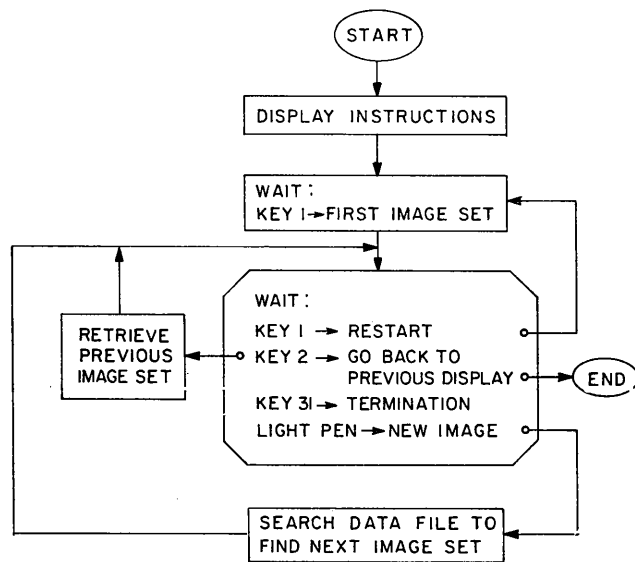


FIG. 5. Flow chart for the program.

LINKAGE OF IMAGES

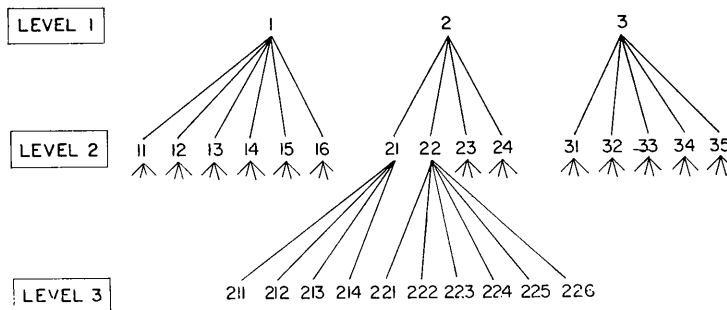


FIG. 6. Diagram of the way in which 'key words' used by the computer in locating appropriate images are built up.

measurements indicated on it) he should now be able to select the appropriate generic image. If he chooses and intercepts the first image of the array, the computer will immediately produce a fourth display (Fig. 4) which will give, in addition to a slightly refined diagram of the organism, its name and a collection of useful data about it. These include: where it was first described, the most recent authoritative publication about it, its diagnostic features, its dimensions, the substrates on which it is found or from which it has been isolated, and the media on which it is commonly cultured. The operator can then check these data against his organism to confirm his identification.

The system is flexible in that, as the operator contemplates each set of images, he has three alternatives: (1) to proceed to the next array, (2) to return to the previous array, or (3) to go back to the beginning of the program; so misgivings and missteps can be quickly checked or corrected. If the operator is an accurate observer and a deft decision-maker, he should, with the aid of the program outlined in this paper, be able to identify his unknown Hyphomycete in less than a minute.

The computer hardware used in this study comprised an IBM 2250 Model 1 graphic display unit on-line with an IBM 360/75 computer.

The software comprised a program written in Fortran IV, and the IBM Graphic support package (GSP) which was used to interface the 2250 with Fortran. Each image was plotted by hand (the 'heroic' method) on a 512×1024 matrix, every point being specified by an X axis and a Y axis address.

An elementary list-structure of key words was used to link the images into sets for each display. The image sets were likewise linked in chains from the generalized shapes to the specific, named images.

The major problems associated with a program such as might be envisaged, involving a very large number of images, are effective mass storage devices and an efficient algorithm for finding the proper images for each display.

We did not consider the storage problem, since our small sample set of images could easily be held in core. Our algorithm examined

key words or pointers to determine which images belonged to the required display. The flow chart (Fig. 5) describes what the program does from the angle of the operator sitting at the 2250 Graphic display unit. The 'keys' referred to are the function keys on the 2250. The light pen contains a photocell which responds to light emanating from the screen (here the 'pen' derives its information from the 'paper,' rather than the reverse). When the operator presses one of the function keys or uses the light pen, the program is interrupted and continues along the particular path defined by this interruption.

The logic of the program itself closely follows that shown in the flow chart. First, data are read and instructions are displayed to instruct a novice in the use of the program. The program goes into a 'wait' condition, that is, it does nothing, until the operator makes an appropriate response. The computer then searches its stored data to find the images in the first display (level 1). In Fig. 6 it is shown how the various images are given key words. After these images have been displayed, the computer again waits for a response from the operator. According to the nature of that response, the computer searches its data file to abstract the new set of images. The new set is generated. Again the program waits, and so on until termination.¹

Acknowledgments

I thank Mr. Martin Ward who both wrote the program for this study, and heroically plotted the images, and Professor Maurice Constant for his encouragement and assistance.

Grateful acknowledgment is made to the National Research Council of Canada for financial assistance.

¹A short black and white 16-mm film has been made of this program being used by an operator at the IBM 2250 console. Copies of this film may be obtained from Prof. Bryce Kendrick, Department of Biology, University of Waterloo.

A write-up of the program has been placed in the Depository of Unpublished Data, NRCC, and copies may be obtained from

The Depository of Unpublished Data,
National Science Library,
National Research Council of Canada,
Ottawa, Canada K1A 0S2.

- VON ARX, J. A. 1970. The genera of fungi sporulating in pure culture. Cramer, Lehre.
- BARRON, G. L. 1968. The genera of hyphomycetes from soil. Williams and Wilkins, Baltimore.
- ELLIS, M. B. 1971. Dematiaceous hyphomycetes. Commonwealth Mycological Institute, Kew.
- KENDRICK, B., and J. W. CARMICHAEL. The Hyphomycetes. *In* The fungi—an advanced treatise. Vol. IV A. Edited by Ainsworth, Sparrow, and Sussman. Academic Press, New York and London. In press, 1972.

